

**IN THE UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK**

HACHETTE BOOK GROUP, INC.,
HARPERCOLLINS PUBLISHERS LLC, JOHN
WILEY & SONS, INC., and PENGUIN
RANDOM HOUSE LLC

Plaintiffs,

v.

INTERNET ARCHIVE and DOES 1 through 5,
inclusive

Defendants.

Case No. 1:20-CV-04160-JGK

**DECLARATION OF ELIZABETH
A. MCNAMARA**

Pursuant to 28 U.S.C. § 1746, I, **ELIZABETH A. MCNAMARA**, declare as follows:

1. I am an attorney duly admitted to practice law before this Court and a partner of Davis Wright Tremaine LLP, counsel for plaintiffs Hachette Book Group, Inc. (“Hachette”), HarperCollins Publishers LLC (“HarperCollins”), John Wiley & Sons, Inc. (“Wiley”) and Penguin Random House (“PRH”) (collectively, the “Publishers”) in the above-captioned action. I submit this declaration in support of the Publishers’ motion for an order granting summary judgment on the Publishers’ copyright infringement claims against defendant Internet Archive (“Internet Archive” or “IA”). Except as otherwise indicated, I make this declaration based on discovery in this case, as referenced by citation, or my personal knowledge.

**INTERNET ARCHIVE OPERATES A WEBSITE THAT INFRINGES IN-COPYRIGHT
BOOKS ON AN INDUSTRIAL SCALE**

1. This action concerns the unlicensed reproduction and distribution of in-copyright ebooks on Internet Archive’s interrelated archive.org and openlibrary.org websites (collectively, the “Website”). (A true and correct copy of the Complaint in this action is annexed as Exhibit 1.) More specifically, this action concerns the “Books to Borrow” portion of the Website to which

Internet Archive directly uploads ebooks. When this lawsuit was commenced, Internet Archive admitted that “more than 1.3 million [in-copyright] books” were available to read in their entirety on the Website. (A true and correct copy of the Answer filed in this action (the “Answer”) is annexed as Exhibit 2.) (*See id.* at ¶3.) Since that time, the number of unlicensed in-copyright ebooks available on the Website’s “Lending Library” has increased to over 3.4 million. (A true and correct copy of the “Books to Borrow” page on archive.org is annexed hereto as Exhibit 3.)

A. IA’s Website

2. The Website effectuates about 70,000 “borrowing events” per day – which amounts to 25 million per year. (A true and correct copy of the Declaration of Brenton Cheng dated January 31, 2022 is annexed hereto as Exhibit 4.) (*See id.* ¶¶40-41.) (A true and correct copy of relevant excerpts from the transcript of the December 9, 2021 deposition of Brewster Kahle (the “Kahle Tr.”) is annexed hereto as Exhibit 5.) (*See id.* at 23:21-23.) Internet Archive lists 5.9 million users, up from 2.6 million when this action was filed. (A true and correct copy of the “users” webpage is annexed hereto as Exhibit 6.)

3. Anybody can “sign up” for free access to the Website “from anywhere in the world” by entering “[b]asic contact information including [an] email address.” (A true and correct copy of relevant excerpts from the transcript of the December 17, 2021 deposition of Chris Freeland (“Freeland Tr.”) is annexed hereto as Exhibit 7.) (*See id.* at 177:7-12.) In other words, IA “[l]end[s] to anyone with an Internet Archive library card, and there is no way to restrict usage of books put into [controlled digital lending] by particular users.” (A true and correct copy of the relevant email from Chris Freeland is annexed hereto as Exhibit 8.) “Controlled digital lending” or “CDL” refers to the theory that Internet Archive and its allies developed to try to argue that it should be lawful for it to scan print books to create ebooks and then loan them out subject to certain limitations,

including lawfully acquiring a print copy for each ebook loan and digital rights management (“DRM”) technology to prevent further reproduction of the ebooks by the user. (A true and correct copy of a July 29, 2020 blog post on archive.org titled “Libraries lend books, and must continue to lend books: Internet Archive responds to publishers’ lawsuit” is annexed hereto as Exhibit 9.) (“Through CDL, libraries lend a digitized version of the physical books they have acquired as long as the physical copy doesn’t circulate and the digital files are protected from redistribution.”)

4. At the time this action was filed and long before, any user who logged in with a valid IA account could “borrow” for free in-copyright ebooks for a period of 14 days. (*See* Answer, ¶77; Kahle Tr. 235:4-9, stating Internet Archive permitted users to “borrow” all books for a 14 day period through the period before this action was filed; July 7, 2022 Declaration of Ian Foster (“Foster Decl.”) at ¶ 45.) After this action was filed, IA changed its practices and now allows valid users to “borrow” free in-copyright ebooks for a period of one hour or 14 days, depending on how many physical copies of the book Internet Archive has in its possession. (*See* Freeland Tr. 177:18-179:25.) If the lending cap for a particular ebook title has been exceeded, the user goes on a waitlist until it becomes available. (Foster Decl. ¶ 44.) Once an ebook is checked out, a user can “flip through or read all the pages, absolutely.” (Kahle Tr. 89:21-25.) Internet Archive, by its own account, offers “High Quality Page Images” of the books available for download on its Website. (A true and correct copy of a document titled “Lending & Book Reader” is annexed hereto as Exhibit 10.) (*See id.* at INTARC00138770.) Internet Archive’s expert, Dr. Imke Reimers, testified that the quality of the scans was “fine,” “quite similar to the Google Books scans” and sufficiently clear to serve as a “potential substitute” for authorized library or commercial ebooks. (A true and correct copy of relevant excerpts from the transcript of the June

3, 2022 deposition of Dr. Imke C. Reimers (“Reimers Tr.”) is annexed hereto as Exhibit 11.) (*See id.* at. 21:19-23:11.)

5. Users of the IA Website use that platform to read books and Internet Archive encourages them to do so. The top of the homepage of openlibrary.org invites readers to “Read Free Library Books Online.” (A true and correct copy of the openlibrary.org homepage is annexed hereto as Exhibit 12.) An Internet Archive user testified in this action that they “can read the whole book” they check out from the Website and have “sometimes read long enough that I reach the end of the hour limit and had to renew.” (A true and correct copy of relevant excerpts from the transcript of the March 24, 2022 deposition of Laura Gibbs (“Gibbs Tr.”) is annexed hereto as Exhibit 13.) (*See id.* at 44:11-45:3.) When Internet Archive changed its default loan period from 14 days to one hour, it received dozens of complaints from users with complaints like “How come all the books I try to read have a one hour limit? I’m confused and I’d just love to finish a series I’m reading.” (True and correct copies of correspondence from Website users are annexed hereto as Exhibit 14.) There are more than 1.1 million ebooks still listed for 14-day downloads on the Website. (*See* Exhibit 3.)

6. The IA’s FAQs address a scenario where “[t]he book I want to read is only available for a 1 hour loan, but I can’t read the whole book in that time. Can I renew a book that I have borrowed for 1 hour?” The answer is: “Yes. If there are available copies, patrons can check the book out again. Also, if a patron continues reading the book, turning pages, after the 1 hour duration has passed and there is still a copy available, the book will be auto-renewed for another hour.” (A true and correct copy of the archive.org FAQ page titled “Borrowing From The Lending Library” is annexed hereto as Exhibit 15.) And, for the books that are available for one hour loans only, the top of the page view reassures users that the loans are “[r]enewable every hour, pending

availability.” (A true and correct copy of the book view page for *Middle School* by James Patterson and Chris Tebbets is annexed hereto as Exhibit 16.)

B. The Works in Suit

7. The Publishers commenced this copyright infringement action on June 1, 2020, alleging that Internet Archive was infringing 127 of their in-copyright books (the “Works in Suit”), which range from literary classics, like *The Bell Jar* and *Lord of the Flies*, to groundbreaking masterpieces, like Zora Neale Hurston’s *Their Eyes Were Watching God* or Toni Morrison’s *Bluest Eye*, to popular non-fiction and fiction titles, like Malcolm Gladwell’s *Blink*, Gillian Flynn’s *Gone Girl*, Ann Patchett’s *Commonwealth*, and George R.R. Martin’s *A Dance with Dragons* (from the *Game of Thrones* series), to perennial children’s titles, like the *Big Nate* or *Lemony Snicket* books, to biographies, to romance novels, like CJ Redwine’s *Defiance*, and even Patrick Lencioni’s best-selling management books. (A true and correct copy of Exhibit A to the Complaint listing the Works in Suit is annexed hereto as Exhibit 17.) The Works in Suit are a small fraction of the more than 33,000 in-copyright books published by the Publishers that Internet Archive has posted as ebooks on the Website without authorization, license or any compensation. (See Foster Decl. ¶¶ 110-115.)

8. By virtue of their respective publishing agreements with authors, the Publishers collectively hold the exclusive rights to publish the Works in Suit pursuant to 17 U.S.C. § 106. (A true and correct copy of the June 10, 2022 Stipulation Regarding Undisputed Facts is annexed hereto as Exhibit 18.) (See *id.* ¶¶ 3-6.) For the respective Works in Suit, the Publishers obtained from the author the exclusive rights to publish the underlying work in multiple formats, including hardcover, paperback, and ebook formats. For most (if not all) Works in Suit, the Publishers also hold audiobook rights in digital and analogue formats. The Publishers generate revenue – and pay

authors royalties – by exercising their exclusive rights over their books in all formats, no less for ebooks than for print or audio editions. All 127 Works in Suit are available as authorized ebooks, which retail consumers pay to read and libraries pay to license so that they can be loaned for free to their patrons. (*See* Declaration of Ben Sevier (“Sevier Decl.”) ¶9; Declaration of Jeffrey Weber (“Weber Decl.”) ¶20; Declaration of Chantal Restivo-Alessi (“Restivo-Alessi Declaration”) ¶20; Declaration of Alan Pavese (“Pavese Decl.”) ¶18.) Over the three years prior to this action, the Works in Suit collectively generated over \$1 million in library ebook revenues alone for the respective Publishers and their authors. (*See* Restivo-Alessi Decl. ¶49; Weber Decl. ¶23.)

C. Library Ebook Market and IA’s Refusal to License Ebooks

9. Internet Archive’s library expert, Susan Hildreth, testified that there is a “thriving ebook licensing market for libraries” that “has increased in recent years” and “is predicated on licensing revenues that are paid by libraries to entities like OverDrive.” (A true and correct copy of relevant portions of the transcript for the May 17, 2022 deposition of Susan Hildreth (“Hildreth Tr”) is annexed hereto as Exhibit 19.) (*See id.* at 47:11-23.)

10. The “one-copy/one-user” models under which PRH, Hachette, HarperCollins and Wiley license ebooks to libraries charge a fee to provide access to the work regardless of whether patrons ultimately read short excerpt from a book or the whole thing. (*See* Weber Decl. Ex. 21, PRH ebook terms of sale for public libraries, school libraries and special library wholesalers; Restivo-Alessi Decl. Ex. 4, HarperCollins agreement governing distribution of digital content to libraries by OverDrive; Sevier Decl. Ex. 6, Hachette ebooks and digital audiobooks terms of sale for libraries; Pavese Decl. Ex. 5, Wiley agreement governing distribution of ebooks to libraries by OverDrive.) The same is true for commercial licenses – like print books, a purchaser pays for the work regardless of whether they actually read any or all of the book.

11. Library ebook lending is facilitated by aggregators, the largest being OverDrive which licenses authorized ebooks to more than 8,000 public and academic libraries for free distribution to their patrons. (A true and correct copy of relevant portions of the transcript for the January 31, 2022 deposition of Steve Potash (“Potash Tr”) is annexed hereto as Exhibit 20.) (*See id.* at 141:6-13, 161:12-162:3.) The number of licensed library ebooks checked out via OverDrive was [REDACTED] in 2010, but this increased [REDACTED] to nearly [REDACTED] checkouts in 2020; additionally, in 2012, OverDrive processed 70 million total digital checkouts (including both ebooks and audiobooks); by 2020, the number of total digital checkouts ballooned to 430 million. (A true and correct copy of checkout data produced by OverDrive in this action, along with two press releases issued by OverDrive, is annexed hereto as Exhibit 21.) During that same time period, the number of different ebook titles checked out by library patrons via OverDrive increased from [REDACTED] to more than [REDACTED] – reflecting the profusion of choice as book publishers committed more books to library ebook lending. (A true and correct copy of book data produced by OverDrive in this action is annexed hereto as Exhibit 22.) And total library expenditure on electronic materials (which incorporates digital text materials like ebooks) increased from \$373 million in 2017 to more than \$440 million in 2019, which is a measure of the established and growing demand for these products and a source of vital income for authors. (True and correct copies of spreadsheets compiled by the Institute of Museum and Library Studies reflecting this data are attached hereto as Exhibit 23.) The Publishers are all active in the library ebook market, as detailed in their declarations for summary judgment in this action. (*See Restivo-Alessi Decl.* ¶28; *Sevier* ¶101; *Weber Decl.* ¶6; *Pavese Decl.* ¶3.)

12. Unlike the thousands of libraries that pay the Publishers for authorized ebooks, Internet Archive does not “license materials” for its website, apparently as a matter of principle.

(*See* Freeland Tr. 96:14-15; Kahle Tr. 80:20-85:23.) Rather, to the highly limited degree that it buys ebooks (instead of creating its own by scanning paper books), Internet Archive insists on “purchasing” ebooks “in the clear” – *i.e.*, buying “non-encrypted” files outright without any contractual preconditions on their further use. (Kahle Tr. 191:21-195:1.) Internet Archive has identified only a handful of publishers that have agreed to sell ebook files on these terms – the literary press 11:11 and the affiliated anarchist publishers PM Press and AK Press. (*See* Kahle Tr. 139:23-140:12.)

13. In the absence of permission to distribute the ebooks on its Website – and despite thousands of demands from copyright owners to remove their in-copyright books (*see* ¶¶16-26, *infra*) – Internet Archive scans physical books in bulk and “republishes” unauthorized ebook editions on the Website. (A true and correct copy of relevant portions of the transcript for the December 3, 2011 deposition of Brenton Cheng (“Cheng Tr.”) is annexed hereto as Exhibit 24.) (*See id.* at 59:18-60:6.) With respect to the Works in Suit, Internet Archive admits that it “digitally scanned physical books embodying the Works” and “provided one or more logged-in patrons with one or more digital versions of the physical books embodying the Works through the lending functions of [the Website].” (A true and correct copy of Defendant Internet Archive’s Objections and Responses to Plaintiffs’ First Set of Requests for Admission is annexed hereto as Exhibit 25.) (*See id.* at 5-6.) Internet Archive has admitted that it lacked permission from rightsholders to engage in any of these activities. (*See* Kahle Tr. 66:22-67:5; 69:8-14.)

14. Internet Archive’s conduct will deprive the Publishers (and their authors) of revenue. Internet Archive does not pay the license fees that libraries pay for authorized ebooks. And, because IA’s ebooks offer a free substitute for authorized ebooks, the Publishers lose the income they would have received from paying consumers of their authorized ebook products who

use IA ebooks instead. Susan Hildreth, Internet Archive’s library expert, stated in her report that libraries engage in CDL “to leverage their existing physical collection to better serve the reading population.” (A true and correct copy of the February 25, 2022 Expert Report of Susan Hildreth is annexed hereto as Exhibit 26.) (*See id.* at ¶ 11.) She also opined that “if a library decided not to license a title because digitized print copies are available for borrowing under CDL, a library will use the money it would have spent licensing that title as an ebook on licensing another ebook title – or on purchasing print books.” (*Id.*) At her deposition, Hildreth further testified that it is her “expert opinion that libraries will spend less money on licensing the e-book editions of the particular titles that were provided through CDL.” (Hildreth Tr. 226:24-4.) This testimony confirms that the logical outcome of CDL is less revenues for the publisher and author of the particular titles IA distributes for free on its website.

15. Kahle has stated that he sees the copying and distribution of unauthorized ebook versions of in-copyright books as a “rebellious” act that “SIMPLY ASSUME[s]” that IA “can [p]erform a function in the online environment that [libraries] routinely perform in the physical print environment....” (A true and correct copy of a document produced by Internet Archive in this action titled “Everyone Deserves to Learn” is annexed hereto as Exhibit 27.) (*See id.* at INTARC00142135.) Internet Archive continues to grow its Website of free, unlicensed, in-copyright ebooks for anyone in the world with an internet connection (*see* ¶¶ 62-76, 99-109, *infra*), notwithstanding opposition from the Publishers, numerous author organizations (like the Authors Guild), the Association of American Publishers, the U.K. Society of Authors and many individual authors, who have all expressly demanded that IA stop its practices and remove works it does not own or license.

**INTERNET ARCHIVE HAS RECEIVED THOUSANDS OF COMPLAINTS FROM
AUTHORS AND PUBLISHERS OVER ITS WEBSITE**

16. Internet Archive has fielded complaints and concerned inquiries about its practices ever since it started republishing in-copyright books without permission. It has claimed that it responds to takedown requests from rightsholders by removing their content. (A true and correct copy of correspondence between author Mark Stein and Internet Archive’s Jeff Kaplan is annexed hereto as Exhibit 28.) When contacted in May 2016, for instance, by a librarian expressing concerns about contributing in-copyright books to the Website, Kahle replied that “[i]n general, we try to put things up and see if there are any problems. In 99% of cases people are just happy, and if there are any issues we just take the material down.” (A true and correct copy of relevant email correspondence between Kahle and Elliot Wrenn of the United States Holocaust Memorial Museum is annexed hereto as Exhibit 29.)

17. Notwithstanding Kahle’s representation, each of the Publishers has requested that Internet Archive take down many thousands of its authors’ books – including Works in Suit – and Internet Archive has failed to comply with those requests or prevent the books from reappearing on the Website. (See Foster Decl. ¶¶ 119-35.) Wiley sent Internet Archive a request to take down *Reminiscences of a Stock Operator* by Edwin Lefevre on November 14, 2011. (A true and correct copy of the relevant email from Wiley to Internet Archive is annexed hereto as Exhibit 30.) That book is still available to be downloaded from the Website. (A true and correct copy of the webpage for *Reminiscences of a Stock Operator* on openlibrary.org and the browser view for the checked out book is annexed hereto as Exhibit 31.)

18. In March 2014, counsel for PRH sent Internet Archive emails stating that its “library lending platform” was “infringing our copyrights” and undermining the “vital and important business segment” for authorized library ebooks. (A true and correct copy of the relevant emails from PRH to Internet Archive is annexed hereto as Exhibit 32.) PRH demanded,

as a first step, that Internet Archive remove 29 Random House titles, including *The Bluest Eye* by Toni Morrison. *The Bluest Eye* is a Work in Suit that, despite IA’s assurances that “borrowing access ha[d] been disabled,” was available to download from the Website when this action was commenced on June 1, 2020.

19. PRH followed up with an email dated July 23, 2014 stating that all “[u]nauthorized, scanned versions of Penguin Random House titles being lent to patrons need to removed from the Open Library site immediately.” (A true and correct copy of this email is annexed hereto as Exhibit 33.) A “full list of Penguin Random House titles and isbns” – including several Works in Suit – was sent to Internet Archive to facilitate the takedown and Internet Archive agreed to “process the disablement of lending access for these items,” but ultimately failed to effectively remove PRH titles from the Website. (*Id.*) (A true and correct copy of the relevant spreadsheet is annexed hereto as Exhibit 34.) Recently there were 16,496 PRH titles available on the Website. (*See Foster Decl.* ¶ 115.)

20. On January 11, 2018, counsel for HarperCollins sent Internet Archive a letter demanding that it “immediately disable the ‘Borrow eBook’ function for our books still under copyright.” (A true and correct copy of this email is annexed hereto as Exhibit 35.) There were recently 7,055 HarperCollins books available on the Free Ebook Website. (*See Foster Decl.* ¶ 115.)

21. Since at least 2015, Hachette and its affiliates submitted hundreds of notices listing URLs for Internet Archive to take down from the Website. In 2019, for instance, Hachette (through its anti-piracy vendor) sent multiple requests to remove Malcolm Gladwell’s *Blink* and *What the Dog Saw* – which are both Works in Suit that were still available on the Website when this action commenced in June 2020. (True and correct copies of relevant portions of takedown

requests for Malcolm Gladwell books are annexed hereto as Exhibit 36.) There were recently 4,519 Hachette books available on IA's Website. (*See* Foster Decl. ¶ 115.) In addition to the specific instances listed above, the Publishers and their agents have sent Internet Archive requests to take down hundreds of titles that Internet Archive has failed to effectuate. (*See* Foster Decl. ¶¶ 117-33.)

22. Internet Archive has also received requests from authors to remove many thousands of books from the Website. To select a few examples from many, Internet Archive received correspondence from a lawyer representing two authors who suffered "significant damage" from IA posting an ebook copy of their book because "their publisher has now decided not to move forward with the release of a new edition." (A true and correct copy of this correspondence is annexed hereto as Exhibit 37.) Another author told Internet Archive that she was "spending every waking hour of my life sending take-down notices. I am spending 50 hours a week to send take-down notices to mitigate the damage this has done to my career." (A true and correct copy of this email is annexed hereto as Exhibit 38.) Yet another informed Internet Archive that "[m]any authors (like me) are releasing their out-of-print-titles as ebooks" and "[i]f we lose a sale because you offer our work for free, that is a definite profit issue for us." (A true and correct copy of this email is annexed hereto as Exhibit 39.)

23. In 2011, the son of Ken Kesey contacted Internet Archive to request that his fathers' books be removed from the Website. (A true and correct copy of relevant emails from Zane Kesey to Internet Archive's Office Manager, Chris Butler, are annexed hereto as Exhibit 40.) Despite Butler's assurance that "[t]he digital lending feature ha[s] been disabled for all [Kesey's] titles..." (*Id.*), works of Ken Kesey – including *One Flew Over the Cuckoo's Nest*, *Sometimes a Great*

Notion and *Demon Box* – are all currently available to borrow from the Website. (A true and correct copy of the author page for Ken Kesey on openlibrary.org is annexed hereto as Exhibit 41.)

24. And Internet Archive received an email from an author writing “from a little farm house in Tasmania where [she had] existed on a meager writer income for the past ten years in order to support my two children as a single mother,” who told IA that it was “stealing” from her by posting her books and had imperiled her ability to “feed [her] kids and pay [her] house loan.” (A true and correct copy of this email and Internet Archive’s response is annexed hereto as Exhibit 42.) The author asked “[c]ould you please consider an agreement between myself and my publishers that will offer remuneration for any future lendings. If not could you remove [my] titles from your system.” “Incidentally,” the email continued, “I noticed *Girls Night In 4* was listed... the authors in that collection gave those stories over for free to support the War Child charity. I’m concerned that [a] valuable charity is missing out on sales. Could you also please consider within your heart and soul if what you are doing is honourable [sic] towards authors. We all love books and we all love sharing books, but the original creator needs to be paid at some stage in your system.”

25. In response, Internet Archive did not acknowledge the author’s request for payment. Instead, a stock response was sent by “The Internet Archive Team” stating that “we have made a good faith effort to disable lending access to any ebook instances of the identified works” on the Website. The email further stated that “[w]e seek to be respectful to creators and operate within the traditional norms and functions of libraries” – and directs the author to “[a] statement from legal scholars about th[e] model” under which Internet Archive republished her books without permission. As of June 29, 2022, seven in-copyright books by the author from Tasmania were still available to read on Internet Archive’s Website. (True and correct copies of

the webpage for that author on openlibrary.org and screenshots of checked out pages from her books are annexed hereto as Exhibit 43.)

26. Sandra Cisneros, the author of the acclaimed *The House on Mango Street* and 12 other books of fiction, poetry and essays, testified that she found the experience of seeing her works distributed to anyone for free on IA's Website "so viscerally upsetting that I could not stay on the website for long." (See Declaration of Sandra Cisneros dated July 5, 2022 ("Cisneros Decl.") ¶ 11.) She considers "Internet Archive's distribution of my books to be a terrible violation of the control I have worked so hard to establish over my work." (Cisneros Decl. ¶ 12).

**INTERNET ARCHIVE'S PURSUIT OF "UNIVERSAL ACCESS TO KNOWLEDGE"
WITHOUT PERMISSION FROM RIGHTSHOLDERS**

27. The Website embodies the aspirations expressed in Internet Archive's motto: "Universal Access to All Knowledge." Kahle Tr. 23:24-24:6.

28. The Internet Archive was founded in 1996 by Brewster Kahle, who currently serves as its Chairman. (See Kahle Tr. 23:21-23.) Kahle is a computer scientist who sold two technology companies in the 1990s, one to AOL and another to Amazon. (Kahle Tr. 17:10-12, 19:23-21:2.) The considerable profits from those sales were used, in part, to fund the Kahle/Austin Foundation, which Kahle runs (as President) with his wife and which reported assets worth approximately \$120 million in 2019. (A true and correct copy of the 2019 990-PF Return of the Kahle/Austin Foundation is annexed hereto as Exhibit 44.) (See *id.* at p. 2.)

29. Internet Archive is set up as a 501(c)(3) corporation. (A true and correct copy of relevant excerpts from the transcript of the October 22, 2021 deposition of Jacques Cressaty ("Cressaty Tr.") is annexed hereto as Exhibit 45.) (See *id.* at 96:22-97:9.) Internet Archive is funded primarily from Kahle's wealth: he has provided tens of millions of dollars of funding to support IA's activities and "he would use various ways of channeling these funds," including via

the Kahle/Austin Foundation. (*Id.* at 42:2-9; 62:20-24.) (A true and correct copy of the Internet Archive *Open Libraries* Proposal to the MacArthur Foundation 100&Change program is annexed hereto as Exhibit 46.) (*See id.* at INTARC00151161, “IA . . . has a guaranteed source of philanthropic support to sustain its basic services in perpetuity.”) Government funding made up only 1.8% of Internet Archive’s revenue between 2011 and 2020. (Cressaty Tr. at 179:13-180:1.)

30. The primary focus of the Internet Archive at its founding was to “archiv[e] the Internet itself” by copying and preserving webpages. (A true and correct copy of an “About the Internet Archive” webpage from IA’s website is annexed hereto as Exhibit 47.) (*See id.* at p. 1.) To that end, Internet Archive used software to “crawl” webpages and copy them into a searchable, chronological database that allows internet users to see what webpages looked like on the dates they were captured. That database is known as the Wayback Machine and it is not at issue in this litigation.

31. Kahle wrote in a March 1, 1997 article for *Scientific American* that “the continued reduction in price of data storage, and also data transmission, could lead to interesting applications as all the text of a library, music of a radio station, and video of a video store become cost effective to store and later transmitted [sic] in digital form.” (A true and correct copy of Kahle’s *Scientific American* article is annexed hereto as Exhibit 48.) (*See id.* at INTARC00390088.) Internet Archive engaged in “building a digital library” and, around 2000 to 2001, it started uploading public domain books to its website. (Kahle Tr. 25:2-27:24.)

32. Kahle gave a speech entitled “Universal Access to Knowledge” on August 4, 2006. (A true and correct copy of an article reproducing that speech is annexed hereto as Exhibit 49.) In that speech, Kahle stated that “[w]e could actually make the dream of the Library of Alexandria a

reality – the dream of having it all. The idea of having all published – and I’d even suggest the bulk of unpublished – things be universally accessible.” (*Id.* at PLAINTIFFS0001110.)

33. For Kahle, the “universal access to knowledge” ethos appeared to clash with the limited access afforded by the “snippets” from books made available on Google Books (the book scanning project at issue in *Authors Guild v. Google, Inc.*, 804 F.3d 202, 209-11 (2d Cir. 2015)). In a December 20, 2006 article Kahle was quoted as saying “[t]he whole Google Books Search looks like Amazon’s Search Inside the Book. Let’s go open with these collections. These are beautiful books.” (A true and correct copy of the December 20, 2006 zdnet.com article titled “Grant funds open-source challenge to Google library” is annexed hereto as Exhibit 50.) (*See id.* at p. 2.)

34. Within a year, Internet Archive announced that “together with the Boston Public library and the Woods Hole Library, that it would start scanning out-of-print but in-copyright works to be distributed through a digital interlibrary loan system.” (A true and correct copy of an October 22, 2007 *New York Times* article titled “Libraries shun deals to place books on web” is annexed hereto as Exhibit 51.) (*See id.* at p. 1.) The article noted by way of contrast that “Google scans copyrighted works as well, but it does not allow users to read the full text of those books online, and it allows publishers to opt out of the program.” (*Id.* at p. 3.)

35. Internet Archive later discarded its “out-of-print” limitation for scanning in-copyright books. (A true and correct copy of a February 22, 2011 blog post on archive.org titled “Internet Archive and Library Partners Develop Joint Collection of 80,000+ eBooks To Extend Traditional In-Library Lending Model” is annexed hereto as Exhibit 52.) As its practices have increased, it now scans any in-copyright books it can acquire and posts ebook versions on its Website, with the only (current) restriction being that it claims to not post books published within

the last five years – a “rule” that it intermittently violates. (A true and correct copy of a document setting out internal limitations for Internet Archive’s in-library collection titled “Notes about Book Collections and Availability” it attached hereto as Exhibit 53.) (*See id.* at INTARC00471120, specifying that the publication date of any book in the collection must be “between 1925 (inclusive) and ~5 years ago.”))

36. Internet Archive subsequently announced that its interim goal is to digitize more than 4 million in-copyright books and add them to the Website. (A true and correct copy of a March 14, 2018 blog post on archive.org titled “Let’s Build a Great Digital Library Together...Starting with a Wishlist” is annexed hereto as Exhibit 54.) Eight months after this lawsuit was filed, IA announced that it had added its two millionth in-copyright book to the Website on February 3, 2021. (A true and correct copy of a February 3, 2021 blogpost on archive.org titled “Internet Archive’s Modern Book Collection Now Tops 2 Million Volumes” is annexed hereto as Exhibit 55.) IA currently lists more than 3 million in-copyright books on its Website and is on track to meet its 4 million-book milestone within a year. (*See* Exhibit 3.) Kahle has also stated that he has “a realistic goal of 10 million books – the equivalent of a major university library.” (A true and correct copy of an August 1, 2011 article in *The Guardian* reporting this statement is annexed as Exhibit 56.)

37. None of the ebooks that Internet Archive republishes are licensed and Internet Archive does not “pay authors royalties for making their books available for free” on the Website. (Freeland Tr. 98:12-20.)

**INTERNET ARCHIVE BUILT AN INDUSTRIAL – OFTEN COMMERCIAL --
INFRASTRUCTURE TO MEET ITS GOALS AND STOCK THE WEBSITE WITH
EBOOKS**

38. To meet its goals, Internet Archive has developed and scaled up technology to facilitate its mass scanning operation. It has also established multiple channels for obtaining the

physical books it needs to scan in order to distribute the resulting unauthorized ebooks on its Website.

39. By 2006, Internet Archive had “developed [its] own little book scanner to get the cost per page – if you were to scan inside the United States – down to ten cents.” (Exhibit 49 at PLAINTIFFS0001111.) That machine, known as a Scribe, photographs the pages of books, which are turned manually by an operator raising and lowering a V-shaped pane of glass that keeps the pages flat. (A true and correct copy of an IA webpage advertising the “Table Top” version of the Scribe machine is annexed hereto as Exhibit 57.) The photographs of each page of the book are run through software developed by Internet Archive, also called Scribe, to create ebooks. (A true and correct copy of the of relevant excerpts from the transcript of the October 14, 2021 deposition of Andrea Mills (“Mills Tr.”) is annexed hereto as Exhibit 58.) (*See id.* at 84:20-85:24.) The Scribe machine and software are collectively known as the Scribe system. (*Id.*)

40. Internet Archive subsequently created “digitization centers,” where operators run multiple Scribe machines to scan high volumes of books. Around 2009, Internet Archive partnered with a company called Data Datum to scan books in China, where labor is cheap. (Mills Tr. 130:19-134:23.) IA also operates scanning centers in the United States, United Kingdom and Canada. (*Id.*) In or about 2019, IA opened a “superscanning” center in Cebu, Philippines that receives shipping containers full of books and digitizes them for inclusion on the Website. (Mills Tr., 132:3-7.) [REDACTED]

[REDACTED] (*See id.* at INTARC00421456, describing process of shipping books to Cebu facility.) Using these facilities, Internet Archive currently scans about 3,500 books a day, which amounts to 1.28 million books annually. (*See* Exhibit 55.)

41. Despite the considerable resources available to it from the Kahle/Austin Foundation and related Kahle entities, Internet Archive does not license ebooks and does not pay any royalties to authors or publishers in connection with its digitizing and posting books on its Website for free. (Freeland Tr. 98:10-20.) Instead, it stocks its Website by scanning and digitizing physical books that it obtains primarily via two channels – (1) by partnering with libraries to scan their collections, and (2) acquiring print books.

42. Internet Archive enters into contracts with certain libraries to scan their collections as part of a “commercial” book scanning business, which has generated more than \$35 million of income since 2011. (Cressaty Tr. 168:19-169:3). (A true and correct copy of Internet Archive’s profit and loss statement for 2011-2020 is attached hereto as Exhibit 60.) Internet Archive produced a flier inviting libraries to “Digitize your Collections [a]t one of our Regional Digitization centers or on your Table Top Scribe locally” that promises “[u]nlimited downloads of your items, with perpetual access. Keep and distribute copies!” (A true and correct copy of this flier is annexed hereto as Exhibit 61.) Participating libraries enter into an agreement to provide IA with print books from their collections to scan, which can be done on premises or offsite. (A true and correct copy of Internet Archive’s scanning agreement with Providence Public Library (the “Scanning Agreement”) is annexed hereto as Exhibit 62.) The scanning process keeps the book intact and, once scanned, the physical books return to the library for further circulation. (*Id.* at INTARC00355369.)

43. Under standard terms in the Scanning Agreement, Internet Archive “provide[s] one digital copy of each digitized [book] ... to the Library and will retain additional Digital Copies. Internet Archive will post Digital Copies on the Internet Archive in a newly created sub-collection... The Digital Copies will be freely available from Internet Archive via HTTP, Torrent

or a similar method.” (*Id.*) The terms further state that both the library and Internet Archive “may freely use their Digital Copies ... in any manner” that is “permitted under applicable copyright law,” including “reproducing, displaying, storing, modifying, or distributing the Digital Copies.” (*Id.* at INTARC00355370.) The majority of the books Internet Archive scanned from library collections were in the public domain, but some books are in-copyright. Chris Freeland, Internet Archive’s Director of Open Libraries, first testified that in-copyright books scanned from library collections cannot be checked out on its Website. (Freeland Tr. 296:22-305:9.) But, after he was shown an example of a work scanned from the collection of the Boston Public Library that apparently was available for borrowing, Freeland later testified that in-copyright “books may well be available for lending on archive.org that were obtained via scanning agreement...” (*Id.* at 314:15-316:16) (A true and correct copy of Exhibit 295 from the Freeland deposition, showing the archive.org page for the work, is attached hereto as Exhibit 63.)

44. Internet Archive discovered that it could not meet its goals by scanning books from library collections alone. As Kahle wrote in a July 1, 2019 blog post, “[t]he IA has worked with 500 libraries over the last 15 years to digitize 3.5M books. But based on copyright concerns the selection has often been restricted to [public domain] books.” (A true and correct copy of the July 1, 2019 blog post on archive.org titled “Most 20th Century Books Unavailable to Internet Users – We Can Fix That” is annexed hereto as Exhibit 64.) IA’s former Director of Finance also testified that, by 2016, “our library partners ran out of books that were out of copyright, so pre-1923, and they’re reluctant to give us books that were in copyright.” (Cressaty Tr. 174:24-175:2.)

45. Given the reluctance of libraries to permit the scanning of in-copyright books due to “copyright concerns,” Internet Archive looked elsewhere to acquire physical copies of in-copyright print books to scan. As early as 2011, Internet Archive had “gathered about 500,000

books” from donations. Then, in or around 2012 to 2013, IA downloaded nearly 1 million ebooks from the “pirate site” Library Genesis (also known as “LibGen”), including a copy of *Song of Solomon* by Toni Morrison – which is a Work in Suit in this action. (See Kahle Tr. 208:8-211:25.) (True and correct copies of the webpage for *Song of Solomon* identifying it as a “Library Genesis” book and a June 23, 2021 screenshot of the Library Genesis collections page listing 920,366 items are attached hereto as Exhibit 65.) Internet Archive denies distributing LibGen via the Website, but Kahle did ask at least one publisher whether she would “be ok with our opening up the 3k mit press books from libgen for one-at-a-time lending on openlibrary...” (A true and correct copy of an email from Kahle to Amy Brand of MIT is annexed hereto as Exhibit 66.) Internet Archive also purchased books that its scanning contractor, “Datum Data,” had obtained “through their connections in China.” (Mills Tr. 92:13-23.) Internet Archive’s former Digitization Program Manager testified that she is “not aware of where the books come from.” (Mills Tr. 113:7-14.) There are notorious problems with counterfeit books in China. (A true and correct copy of a January 23, 2018 CNBC article titled “Plagiarism is rampant in China, and its media companies are raking in billions” is attached hereto as Exhibit 67.)

46. Internet Archive has also obtained books by absorbing entire library collections. In 2020, for instance, the defunct Marygrove College donated its entire collection of “70,000 books” and other works. (A true and correct copy of an October 20, 2020 blog post on archive.org announcing the donation is annexed hereto as Exhibit 68.) Those materials were packed for shipping, digitized and resulting ebook copies were posted on the Website. (A true and correct copy of the Marygrove College Library page on archive.org is annexed hereto as Exhibit 69.)

47. Internet Archive also has operated a book sponsorship program that allows users to contribute money towards the purchase and scanning of particular books they want to see added

to the Website. As Internet Archive’s former Director of Finance described it, “[i]f you have a book that you are interested in and interested in in having a digital copy of ..., you can send us money or you can send us a book plus a donation to cover the cost of ... the digitization of that book.” (Cressaty Tr. 167:21-168:7.) In other words, “the donor gives Internet Archive money in excess of the price of the book, and that money covers the purchase of the book and its scanning.” (*Id.*; *see also* Complaint ¶99, showing webpage requesting donation of \$93.13 to digitize a copy of Adam West’s biography “Back to the Batcave.”) (A true and correct copy of an October 24, 2019 blog post on openlibrary.org describing the book sponsorship program is annexed hereto as Exhibit 70.)

48. Internet Archive also has funded its projects with donations of large amounts of bitcoin. For instance, Internet Archive received “\$1M in Bitcoin” from the anonymous donor behind the “Pineapple Fund.” (A true and correct copy of a December 26, 2017 blog post on archive.org announcing the donation is annexed hereto as Exhibit 71.) (*See also* Cressaty Tr. 126:24-127:20.) Internet Archive testified that “there’s no way you can tell” where this money is coming from and the source “could be ISIS, for all you know.” (Cressaty Tr. 114:12-116:18.)

49. The books Internet Archive obtains are typically packed into shipping containers and sent to offshore digitization facilities in Cebu or China, where they are rendered into ebooks by the Scribe system. (Mills Depo Tr. at 131:17-132:7.) The books are then shipped back to facilities operated by IA’s affiliates, where they “are stored in double-stacked shipping containers in ... physical archive facilities in Richmond, CA and Pennsylvania.” (A true and correct copy of a presentation by Chris Freeland titled “How controlled digital lending works for libraries” is annexed hereto as Exhibit 72.) (*See id.* at INTARC00142783.) Unlike the physical books in the collections of public libraries, the physical books Internet Archive “preserves” in shipping

containers are not available to the public. (A true and correct copy of a presentation on Internet Archive’s lending and digitization programs is attached hereto as Exhibit 73.) (*See id.* at INTARC00142733; Mills Tr. 176:2-8.)

50. The books and the facilities housing the shipping containers are legally owned by a separate entity, Open Library of Richmond Inc. (“OLR”). (A true and correct copy of OLR’s 2019 990-PF Return is annexed hereto as Exhibit 74.) (A true and correct copy of IA’s 2016 990-PF Return is attached hereto as Exhibit 75.) (*See id.*, reflecting that Open Library of Richmond paid Internet Archive a \$4,840,000 grant; *see also* Kahle Tr. 44:23-45:2, “[T]he Open Library of Richmond [] is actually the organization that stores and owns the books.”; Cressaty Tr. 62:11-14). Kahle is the principal officer of OLR and he funds that entity from his personal finances. (Cressaty Tr. 42:10-14.) Internet Archive has a number of other affiliate companies in addition to OLR. (Cressaty Tr. 37:3-47:4, 78:24-81:10, 60:1-61:1, referring to the companies as “Kahle/Austin Empire.”) These companies – which own IA’s “headquarters, warehouses and some distributed data centers” – were created “[a]s a strategy to mitigate risk.” (*See* Exhibit 46 at INTARC00151161.)

51. In or around 2013, Internet Archive started sourcing physical books from the “socially conscious for profit” company Better World Books (“BWB”), which is a used book retailer that has “reused or recycled” nearly 400 million print books, including from libraries discarding (or “weeding”) unwanted copies. (A true and correct copy of BWB’s homepage listing the number of “books reused or recycled” is annexed hereto as Exhibit 76) (*See id.* at p. 2.) (A true and correct copy of relevant excerpts from the transcript of the December 6, 2021 deposition of Ginger Patton-Schmitt (“Patton-Schmitt Tr.”) is attached hereto as Exhibit 77.) (*See id.* at 85:18-86:10, confirming BWB’s for-profit status.) (A true and correct copy of BWB webpage titled

“About Better World Books,” describing BWB as a “for-profit social enterprise,” is attached hereto as Exhibit 78.) In April 2013, Internet Archive and BWB drafted a memorandum of understanding stating that [REDACTED]

[REDACTED] (A true and correct copy of the draft memorandum of understanding is annexed hereto as Exhibit 79.)

52. During this time period, Internet Archive periodically identified print books that it wanted to acquire based on what was available from BWB’s inventory, which BWB either donated or sold to IA for a modest fee. In an email from Brewster Kahle to BWB’s Director, Strategic Sales & Partnerships, for instance, Kahle stated that IA had identified “about 13k books that were ... ‘desirable’ for us” from BWB’s available inventory and offered to purchase those “books for \$1 each.” (A true and correct copy of this email is annexed hereto as Exhibit 80.)

53. Eventually a system developed whereby Internet Archive sent a “wish list” containing approximately “1.5M ISBNs” for BWB to run against its inventory to determine which books were available. (A true and correct copy of relevant emails between Freeland and a BWB Project Manager is annexed hereto as Exhibit 81.) Kahle and Freeland “agreed that our best path to millions of books is arm-in-arm with BWB...” (*Id.* at INTARC00414924.) By April 2018, IA was [REDACTED] (A true and correct copy of a memorandum entitled [REDACTED] is annexed hereto as Exhibit 82.)

54. In the fall of 2018, Internet Archive started exploring the possibility of acquiring Better World Books. In an email sent to Kahle on November 12, 2018, the founder of BWB wrote that [REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED] (A true and correct copy of emails between Kahle and Xavier Helgesen is annexed hereto as Exhibit 83.)

55. On April 17, 2020, Kahle wrote to a potential investor that Better World Books “get[s] 30 million books a year, and sell[s] 10 million – the others are donated or recycled. We want more of the flow... This is a bold move and our library colleagues believe it can help rearrange how books work in the library field in general: libraries buy books, and now consign or donate them to BWB, the new non-profit would keep 1 copy for digitization and preservation, the rest get sold. This is what BWB does already, but we would ramp this up. As a self sustaining business, this will keep the books flowing year after year.” (A true and correct copy of emails between Kahle and Peter Mounce are annexed hereto as Exhibit 84.) Kahle also wrote that “they are a going concern and hopefully even profitable. If they generate money then they could pay for the digitization...” (*Id.*; *see also id.*, “The successful operation of BWB will provide funding back to The Internet Archive to ensure that it can continue to deliver free services to the world...”) And, Kahle wrote, Internet Archive would obtain “maybe 7 million [books] over the next few years” (*id.*) at a rate of “1mm books per year.” (A true and correct copy of a relevant email Kahle sent on May 6, 2019 is annexed hereto as Exhibit 85.)

56. In June 2019, Internet Archive effectively acquired Better World Books in return for [REDACTED] (A true and correct copy of [REDACTED] is attached hereto as Exhibit 86.) (*See also* Patton-Schmidt Tr. at 70:10-15, [REDACTED])

[REDACTED]

[REDACTED]

[REDACTED]). Kahle used funds from his personal fortune to fund the transaction, which were channeled through OLR. (Cressaty Tr. 148:7-25.) As part of the deal, [REDACTED]

[REDACTED]

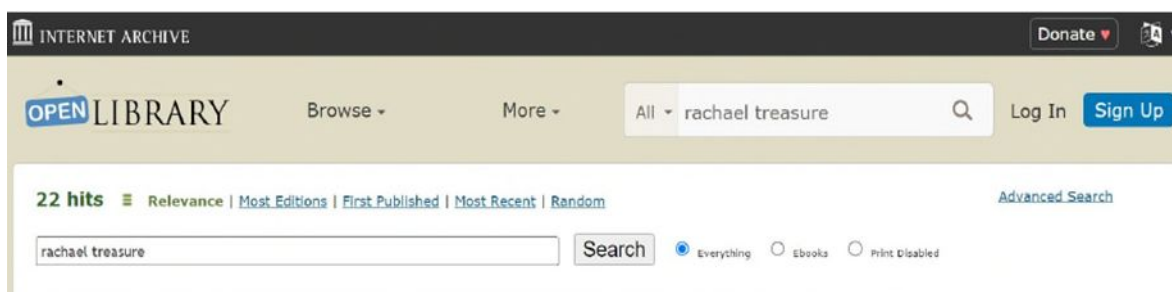
[REDACTED] (A true and correct copy of the July 10, 2019 Donation Agreement between Better World Books and OLR is annexed hereto as Exhibit 87.) Under the terms of the acquisition, Better World Libraries – a 501(c)(3) shell corporation controlled by Kahle – became the sole shareholder of Better World Books. (A true and correct copy of Better World Libraries’ 2019 990-PF Return is attached hereto as Exhibit 88.) (*See id.* at INTARC00402986.) Kahle is the Chairman of Better World Libraries and a Director of Better World Books, which is controlled by “people affiliated with the Austin Kahle [sic] Empire.” (Cressaty Tr. 241:7-16; 248:14-249:10.)

57. Internet Archive announced the acquisition on November 6, 2019 in a blog post that stated that “[t]his new relationship will allow Better World Books to provide a steady stream of books to be digitized by the Internet Archive, thereby growing its digital holdings to millions of books... Any book that does not yet exist in digital form will go into a pipeline for further digitization, preservation and access.” (A true and correct copy of the November 6, 2019 blog post is annexed hereto as Exhibit 89.)

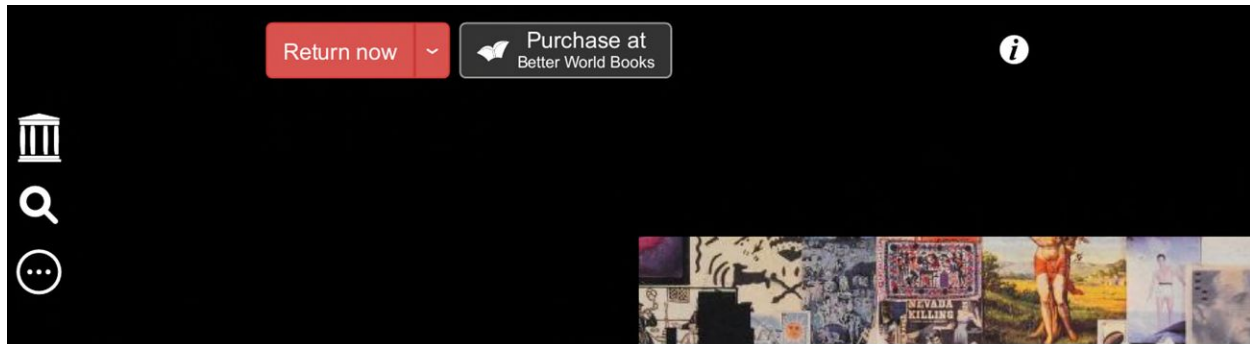
58. Internet Archive and Better World Books have explored and adopted a number of “synergies” – *i.e.*, “ways that BWB could work with IA and its affiliates to make both business perform better.” (Patton-Schmitt Tr. 119:19-23) (A true and correct copy of a document entitled [REDACTED] is annexed hereto as Exhibit 90.) For instance, webpages for books on the Better World Books website were provided with links to “borrow” those books from

the Internet Archive’s Website. (A true and correct copy of an illustrative book webpage on Better World Books’ website is annexed hereto as Exhibit 91.) Better World Books also links to IA ebooks on the Website for its “Preview” feature. (A true and correct copy of the Better World Books preview page for *Americanah* by Chimamanda Ngozi Adichie is annexed hereto as Exhibit 92.)

59. Webpages for ebooks on the Internet Archive’s Website also contain links to purchase that title from Better World Books – including a “Purchase at Better World Books” button that appears at the top of the window when users read an ebook on Internet Archive’s ebook browser. (True and correct copies of documents illustrating the Website’s links to Better World Books are annexed hereto as Exhibit 93.) Internet Archive receives a payment from Better World Books every time a user clicks on a link to purchase a used book from Better World Books’ website. (See Patton-Schmitt Tr. 142:25-154:5.) The webpages on Internet Archive’s Website also include links to “Donate” to the Internet Archive. See Exhibit 43. As Internet Archive’s former Director of Finance testified, “every single page of the Archive is monetized.” Cressaty Tr. 202:14-208:4.



(Excerpt of Exhibit 43.)



(Excerpt of Exhibit 93.)

60. Post-merger, Better World Books [REDACTED] [REDACTED] (Patton-Schmitt Tr. 198:2-9.) “Better World Books has vast management experience in libraries e-commerce and supply chain management that can be beneficial to IA.” (*Id.* at Tr. 120:23-121:8.) Between 2020 and 2021, [REDACTED] [REDACTED] [REDACTED] (A true and correct copy of a Wikipedia page summarizing BWB’s book donations to IA is annexed hereto as Exhibit 94.) (A true and correct copy of [REDACTED] during this time period is annexed hereto as Exhibit 95.)

61. Between the merger in July 2019 and June 30, 2021 Better World Books identified and packed up approximately 2.8 million books from Internet Archive’s wishlist, many of which have been shipped off to the Cebu scanning center for digitization, posting on the Website and (ultimately) delivery to the facilities operated by Internet Archive affiliates for “preservation.” (True and correct copies of dashboard entries showing Better World Books fulfilment for Internet Archive during this time period are annexed hereto as Exhibit 95.) (*See also* Patton-Schmitt Tr. 155:1-11.)

**INTERNET ARCHIVE’S EVOLUTION FROM “IN-LIBRARY LENDING” TO THE
“OPEN LIBRARIES PROJECT”**

62. As previously noted, at present anybody in the world with a valid email account can obtain an account to access any of the millions of in-copyright books available on the Website. Internet Archive has unilaterally placed certain limits on how the ebooks on their website circulate, which have changed over time and could change again.

63. In 2011, Internet Archive had “80,000+” in-copyright ebooks in its “In-Library eBook Lending Program.” (A true and correct copy of the February 22, 2011 blog post on archive.org announcing the launch of the In-Library Ebook Lending Program is annexed hereto as Exhibit 96.) Under this system, each ebook scan could “only be borrowed by one person at a time” for two weeks and patrons could “borrow up to 5 eBooks at a time.” (*Id.*) Ebook scans could “only be borrowed by a patron of a physical library that participates in [Internet Archive’s] program” and the ebooks must be “access[ed] through our site from the physical library’s network.” (A true and correct copy of email correspondence between Chris Butler, Internet Archive Office Manager, and Zane Kesey is annexed hereto as Exhibit 97.)

64. By 2013, Internet Archive had relaxed the geographical limits so that its ebooks could “be borrowed by those in an area served by a physical library which participates in our program.” (A true and correct copy of a February 2013 email from Chris Butler to James D. Jenkins is annexed hereto as Exhibit 98.) (*See id.* at INTARC00165238) (A true and correct copy of a 2012 email from a Library of Michigan librarian is attached hereto as Exhibit 99.) (*See id.*, reporting concern that “[i]t seems like anyone in Michigan can access copyrighted material available through your Open Library In-Library Lending program without authentication.”) Within approximately one year, Internet Archive completely removed geographical limits so that anybody in the world with an internet connection could access in-copyright books on the Website – which remains true to this day. (Kahle Tr. 97:11-18.)

65. The number of concurrent ebook loans permitted has also shifted over time. For several years, Internet Archive loaned books “one at a time” – which meant that it set a cap on the number of users who could simultaneously check out a book based on the number of books Internet Archive itself owned. (Foster Decl. ¶70) So if there were five physical copies of a book in an OLR shipping container, five users could read the ebook version on the Website and any subsequent users would be offered a place on the waitlist. On November 13, 2018, Freeland published a blog post stating that the “own one, loan one principle” based on the books Internet Archive owned was “viable, but limited,” and “for controlled digital lending to work at scale, more physical copies are needed to loan against, especially for titles ... that enter the public zeitgeist and become part of a major news story.” (A true and correct copy of the November 13, 2018 blog post on archive.org is annexed hereto as Exhibit 100.)

66. Kahle explained the idea in broad strokes in a statement entitled “Transforming Our Libraries into Digital Libraries: A digital book for every physical book in our libraries.” (A true and correct copy of this statement is annexed hereto as Exhibit 101.) In that document, he proposed a “collaborative effort [with libraries] to select and digitize the most useful books of the 20th and 21st centuries, and to build a robust system to circulate the resulting e-books to millions, and eventually billions of people.” (*Id.* at PLAINTIFFS0000858.) In order to do this, he proposed collaborating with libraries “to digitize the materials efficiently, minimizing duplication, and lend the digital texts with the same limitations placed on physical books.” (*Id.* at PLAINTIFFS0000860.) Kahle wrote that “Internet Archive could create a circulation system that would administer the lending [for libraries]. In effect, then, each library can choose from a variety of methods to lend digital versions of the physical books in their collection. This would keep the local libraries in control but leverage the convenience of a cloud-based system that others maintain

and update.” (*Id.* at PLAINTIFFS0000865.) To accomplish this, “in each library’s online card catalog, when a digital version of a book exists [on Internet Archive], we can include a web-link on the record for the physical book, giving readers the ability to browse the book on screen or to borrow it from the convenience of their homes. In this way, we can smoothly enhance a library’s collection, from analog to digital, at scale. . . . To build this future, we will need the participation of multiple sectors to bring thousands of libraries digital.” (*Id.* at PLAINTIFFS0000859.) This “collaborative digital library collection and circulations system,” Kahle hypothesized, “could help deliver e-books to millions of patrons with a flip of a digital switch.” (*Id.* at PLAINTIFFS0000865-86.)

67. This concept was embodied in the “Open Libraries” project that Internet Archive promoted in 2018 as a solution to the “limitations of scale for titles with wide appeal” that Internet Archive was experiencing. (Exhibit 100.) This project essentially allows libraries to “pool[] their physical collections” with Internet Archive “in order to make more lendable copies of digital books available to their users and the world.” (*Id.*; *see also* Freeland Tr. 106:1-4.) The purpose of the Open Libraries project is “to increase lending counts” on the Website by “identify[ing] the overlap in [the library’s] physical holdings with our digital holdings and provide free digital books to patrons where there are matches.” (A true and correct copy of the “Join Open Libraries” webpage on openlibraries.online is annexed hereto as Exhibit 102.)

68. Under this system Internet Archive is a central hub and libraries participating in the Open Libraries project – each known as a “Partner Library” – send their catalogues to Internet Archive to “run an overlap analysis that compares [ISBN numbers for] their physical holdings with our digital holdings.” (Freeland Tr. 51:19-24.) Internet Archive does not make new scans of the libraries’ books identified by the overlap analysis (nor take possession of the matching print

editions in the libraries' collections). (*Id.* at 64:6-19.) Rather, every time a book in the libraries' catalogue matches an ebook on the Website, Internet Archive "just increases by one the number of concurrent checkouts of that book" permitted on the Website. (*Id.*) Since Internet Archive does not "set any upper limit to the number of copies available via concurrent lending," the overlap analysis dictates that "if a hundred partner libraries possessed a copy of the same book, the Internet Archive would be able to lend a hundred copies of that book simultaneously." (*Id.* at 65:2-15.) And if there were "a thousand libraries that had the same book and put them into controlled digital lending," then one thousand users would be able to view the ebook scan on the Website." (*Id.*)

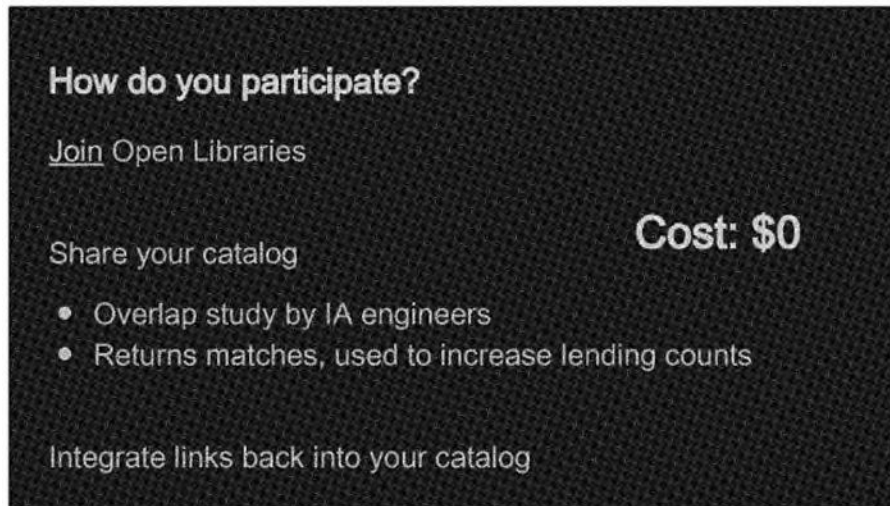
69. In order to become a Partner Library in the Open Libraries project, libraries need only submit a short "online form." (*See* Exhibit 102.) That form forthrightly admits that "Internet Archive's Open Libraries project offers the prospect of making every library's collection into a digital collection by allowing a library to lend a digital version of physical volumes they own." (A true and correct copy of the "Agreement to participate in Open Libraries" (the "Open Libraries Form Agreement") is annexed hereto as Exhibit 103.) Partner Libraries agree to "share their catalog of books with the number of copies of each book with the Internet Archive." The Open Libraries Form Agreement imposes no substantive restrictions on the Partner Libraries' use of the physical books that match the ebooks on the Website. (*Id.*) The Open Libraries Form Agreement further provides that any partner "Library may integrate links to the borrowable Books in their catalog and other services. The Internet Archive will add books to the collections offered on the Internet Archive's sites." (*Id.*)

70. While it is currently free for libraries to become Partner Libraries, Kahle has stated that "it would be understandable if we charged libraries that did not contribute digitization or backend services for access to digital books. It would be equally understandable to charge a one-

time transfer fee to libraries that wanted to store their own local copies” of the ebooks on the IA Website. (Exhibit 101.)

71. Internet Archive has marketed the Open Libraries project and related digitization projects to libraries as a substitute for paying for authorized ebooks via authorized library ebook aggregators and as a means of obtaining “free ebooks for your patrons” with “no cost involved.” (A true and correct copy of a relevant email from Freeland to a librarian at Arizona State University is annexed hereto as Exhibit 104.) For instance:

- a. Freeland sent an email to the Associate Dean of University of Oklahoma Libraries stating that “[t]he main offer we have in hand with the Open Libraries today is the ability to match your physical holdings with the 1.1M in-copyright books we have *already* digitized and where there’s a match, provide you back a link to the digitized book.” (True and correct copies of emails making this offer are annexed hereto as Exhibit 105.)
- b. An Internet Archive representative sent a librarian an email stating that the Open Libraries project “can leverage controlled digital lending to provide your patrons with free ebooks of your physical collections. As an Open Libraries member, we match your in-copyright holdings with our digital holdings and serve free ebooks where they overlap. Join Open Libraries, access free ebooks – so simple, we had to share!” (A true and correct copy of relevant emails from Internet Archive Senior Digitization Manager Jeff Sharpe to Karl Stutzman of Anabaptist Mennonite Biblical Seminary is annexed hereto as Exhibit 106.)
- c. Freeland gave a presentation entitled “Open Libraries Introduction & Internet Archive programs” that included the following slide:



(A true and correct copy of the presentation entitled “Open Libraries Introduction & Internet Archive programs” is annexed hereto as Exhibit 107.)

- d. The notes to a slide for a separate presentation Freeland gave to libraries stated that the Open Libraries project “ensures that a library will not have to buy the same content over and over, simply because of a change in format.” (A true and correct copy of the presentation entitled “Addressing the 20th century gap: Controlled digital lending for in-copyright material” is annexed hereto as Exhibit 108.)
- e. Another presentation was titled “Maximizing institutional investments in print resources through controlled digital lending” – with the subtitle, “**Or, You Don’t Have to Buy it Again!**” (A true and correct copy of the presentation entitled “Maximizing institutional investments in print resources through controlled digital lending” is annexed hereto as Exhibit 109.)
- f. A vendor acting on behalf of Internet Archive sent an email to one of the subscribers to its Library.Link Network, the Evergreen Indiana Library Consortium, with a subject line of “Free eBooks for Indiana Evergreen.” (A true and correct email dated January 3, 2019 from Lauri McIntosh is annexed hereto as Exhibit 110.) That

email stated that “Open Libraries provides a way for you to offer digitized books to your patrons for free” and that joining as a partner would “bring this added value to your libraries.” (*Id.*) (True and correct copies of additional emails from Lauri McIntosh asking subscribers if they “would like to participate and get free ebooks for your end users” are annexed hereto as Exhibit 111.)

- g. The same vendor sent a librarian an email inviting them to join the Open Libraries project, which stated that the “1, 2, 3 of it is, once you sign the form we do the rest... would you like to participate and get free ebooks for your end users?... Internet Archive now offers FREE ebooks for all Library.Link libraries when you participate in the Open Libraries Controlled Digital Lending (CDL) project.” (True and correct copies of emails making this offer are annexed hereto as Exhibit 112.)
- h. Internet Archive’s Website contains a video of a July 14, 2021 presentation entitled “Implementation & Integration: CDL for All Libraries” that describes controlled digital lending as a “[l]ow risk, reasonable solution that preserves legal and fiscal value in collections” and contains the following slide:

How it works:

- Give us your MARC records
- We'll run an overlap with our digital holdings
- We'll give you back links to matched books
- You can bring those links back into your catalog

Cost: \$0

(A true and correct copy of the “Implementation & Integration” presentation is annexed hereto as Exhibit 113.)

- i. Internet Archive gave a presentation at the 2019 Library Leaders Forum that summed up the “[s]teps to participate in Open Libraries: Join Open Libraries, Share your catalog, Overlap study, Integrate links back into your catalog, Lend digital books to your patrons.” (*See* Exhibit 27 at INTARC00142079.)

72. Internet Archive has testified that it has 81 Partner Libraries as of December 2021 (Freeland Tr. 71:4-10) – the majority of which are academic libraries and a small handful of which are public libraries. (As of December 24, 2021, 62 Partner Libraries contributed books to the overlap analysis and 13 of those were public libraries. *See* a true and correct copy of Defendant Internet Archive’s Objections and Responses to Plaintiffs’ Second Set of Interrogatories annexed hereto as Exhibit 114, Response to Interrogatory No. 12.) At the 2020 Library Leaders summit, Kahle announced that “2.8M copies” of in-copyright books were added to concurrent lending counts through the Open Libraries’ project. (A true and correct copy of the “Library Leaders Forum 2020 Partner Summit” presentation is annexed hereto as Exhibit 115.) (*See also* Freeland Tr. 71:1-12.) The number of books added through overlap analysis has increased since that time. (Freeland Tr. 71:14-25.)

73. Many libraries – both Partner Libraries and unaffiliated libraries – have accepted Internet Archive’s offer to incorporate links to ebooks on the IA Website onto their library catalog websites. For instance, the Georgetown University Law Library’s catalog features the language “Full text availability” and presents a link to access the “full text” of the title through “Internet Archive Controlled Digital Lending.” (A true and correct copy of a screen capture from the Georgetown Law Library catalog is annexed hereto as Exhibit 116.) Additionally, the Dartmouth College Library’s catalog entry for “The Catcher in the Rye,” which is one of the Works in Suit published by Hachette, features a link to access an ebook version through the Internet Archive’s

website. (A true and correct copy of a screen capture from the Dartmouth College Library catalog is annexed hereto as Exhibit 117.)

74. Many other unaffiliated libraries have accepted Internet Archive's offer to incorporate links to ebooks on the IA Website onto their main library websites, including public library systems in New York, California, and elsewhere. (A true and correct copy of screen captures from the Burlingame, CA Public Library, the Denver, CO Public Library, the Mesa County, CO Public Library, and Onondaga County, NY Public Library, is annexed hereto as Exhibit 118.)

75. Internet Archive benefits from the overlap analysis by substantially increasing the number of users who can borrow its unlicensed ebooks at one time. For each Partner Library that contributes its catalogue for overlap analysis, Internet Archive gets "an additional copy to lend" of every matching book "without ... having to incur the cost associated with scanning." (Freeland Tr. 66:8-67:14.) "As a result of the Internet Archive implementing overlap analysis ... wait lists [have] been reduced on" the Website "for certain titles." (*Id.*) And the overlap analysis expands the Website's collection in tandem with the book pipeline from Better World Books because "the more books Internet Archive has been able to obtain and scan, the greater likelihood there would be matches in the overlap analysis with potential partner libraries." (*Id.*)

76. As detailed below, in order to induce libraries to become Partner Libraries in the Open Libraries project, the Internet Archive made various statements suggesting that the whole scheme was legal under the Copyright Act.

INTERNET ARCHIVE SPEARHEADS THE CREATION OF THE "CONTROLLED DIGITAL LENDING" THEORY

77. Internet Archive did not publish a formal legal theory to justify its scanning and republication of in-copyright books without permission until 2017. As Kahle stated in the

“Universal Access to Knowledge” speech he gave in 2006, Internet Archive had “found that if you put things out there in a nonprofit setting, it works for people in the sense that they don’t gripe. The idea of opt-out as opposed to opt-in – putting it up and then if somebody complains, taking it down – works very well in these sorts of communities. I would suggest being a bit bold and making things available...” (Exhibit 49 at PLAINTIFFS0001113.)

78. By 2015, the Copyright Office had issued a report concluding that “[t]here is broad agreement that no colorable fair use claim exists” for “providing digital access to copyrighted works in their entirety.” (A true and correct copy of U.S. Copyright Office, “Orphan Works And Mass Digitization: A Report Of The Register Of Copyrights 101” (2015) is annexed hereto as Exhibit 119.) The previous year a subcommittee of the Judiciary Committee of the House of Representatives held hearings about the first sale doctrine, including on whether to adopt a digital first sale doctrine, with strong opposition to a digital first sale doctrine from the Association of American Publishers and others. (A true and accurate copy of the hearing transcript is annexed hereto as Exhibit 120.) While the subcommittee announced reforms for the Copyright Office and Congress passed the Orrin G. Hatch-Bob Goodlatte Music Modernization Act, the Copyright Act was never amended to recognize a digital first sale defense.

79. In 2017, Internet Archive was in the running for a \$100 million grant from the MacArthur Foundation, which Internet Archive intended to use to “turn 80% of library collections digital by 2023.” (A true and correct copy of relevant excerpts from the transcript of the October 18 and 19, 2021 deposition of Lila Bailey (the “Bailey Tr.”) is annexed hereto as Exhibit 121.) (*See id.* at 130-38; *see also* Exhibit 46 at INTARC00151089.) Internet Archive’s in-house policy counsel testified that “as part of the grant proposal, barriers to solving the problem are identified. One of the significant barriers to libraries being able to loan out their digital collections was ...

copyright uncertainty, lack of clarity in the law.” (Bailey Tr. 130:15-19.) The need to articulate a legal theory behind IA’s practices in the application for the McArthur grant is “what precipitated having a meeting in May of 2017 concerning Controlled Digital Lending.” (*Id.* at 130:3-8.) Internet Archive convened the “Open Libraries Copyright Workshop” on May 23-24 to “consider the digitize and lend model that the Internet Archive is proposing to expand through the MacArthur Foundation’s 100%Change grant.” (Exhibit 46 at INTARC00151184.) The Workshop was led by one of IA’s advisors for the grant, the copyright scholar Pamela Samuelson, and was “modeled after a meeting that she did around the Google books case.” (Bailey Tr. 129:19-22.)

80. As per Bailey, Samuelson “wanted to basically have an open discussion, an academic discussion with other scholars and library practitioners who understood how libraries were engaging in this practice.” (Bailey Tr. 130:22-25.) The participant roster for the Workshop submitted as part of the MacArthur Foundation Application included Bailey, Kahle, two other Internet Archive employees, two of the lawyers that have appeared for IA in this action (Joseph Gratz and Corynne McSherry) and several other individuals who would be enlisted to formulate a lending theory – including David Hansen, Kyle Courtney, Jason Schultz, Mary Minow and Michelle Wu. (*See* Exhibit 46 at INTARC00151185.) Bailey admitted that she was not aware of “any copyright lawyers who represent copyright creators at this panel.” (Bailey Tr. 134:18-135:3.)

81. At a dinner shortly after the Workshop, Bailey “started talking about the idea of writing something up” with Courtney, Hansen and Wu. (*Id.* at 138:1-19.) Michelle Wu was a Georgetown law professor and librarian who was another advisor to Internet Archive for the McArthur Grant and, according to Internet Archive, “crafted the legal theory behind Controlled Digital Lending.” (A true and correct copy of an October 20, 2020 blogpost announcing that Wu was awarded the Internet Hero Award for Establishing the Legal Basis for Controlled Digital

Lending is annexed hereto as Exhibit 122.) Wu also was retained by Internet Archive as its counsel, but did not reveal that affiliation in her letter submitted in support of IA's MacArthur grant application. (See Exhibit 46 at INTARC0015123). Kyle Courtney is a Copyright Advisor at Harvard University and David Hansen was the Lead for Copyright and Information Policy and Associate University Librarian at Duke University. (A true and correct copy of a white paper on controlled digital lending by Courtney and Hansen (the "White Paper") is annexed hereto as Exhibit 123.) Bailey has described Wu, Courtney and Hansen as "[f]riends of the Internet Archive." (Bailey Tr. 115:6-10.) At their dinner, Bailey and the "friends of the Archive" decided to draft "something short and easy for a non-lawyer to understand about the legal underpinning of Controlled Digital Lending." (Bailey Tr. at 138:1-19.) They subsequently decided to also publish a white paper that would address the legal issues more fully. (*Id.* at 138:5-10.)

82. The MacArthur Foundation rejected the Internet Archive's application for the \$100 million grant, but the project to draft a statement defining controlled digital lending (the "Statement") continued. The core drafters of the Statement were IA's in-house counsel (Bailey), Courtney, Hansen, Wu, Mary Minow and Jason Schultz, a law professor at NYU. (A true and correct copy of an email from Kyle Courtney to Wu, Bailey, Minow, Schultz and Hansen attaching a draft of the White Paper that "addressed nearly all of your comments, edits, suggestions, and grammatical insights" is annexed hereto as Exhibit 124) Courtney and Hansen were also subsequently selected to draft the White Paper, which Bailey edited prior to publication. On July 12, 2018, for instance, Bailey sent Courtney and Hansen an email with "higher level and structural thoughts," including revising the analysis of the first factor from stating that CDL is not transformative to indicate "that at least some stakeholders in the library community think CDL is transformative." (A true and correct copy of this email is annexed hereto as Exhibit 125.) Bailey

also asked Courtney and Hansen to revise the fourth factor to “hit really hard on why the 6 controls are why we win on the market harm analysis” and emphasize cases under the third factor where wholesale copying was held to be fair use. (*Id.* at INTARC466830.) Bailey provided a further round of edits on August 3, 2018. (A true and correct copy of Bailey’s August 3, 2018 email is annexed hereto as Exhibit 126.)

83. When asked about the status of the White Paper, Freeland stated that “IA is officially NOT leading the effort for the sake of being impartial.” (A true and correct copy of Freeland’s August 20, 2018 email is annexed hereto as Exhibit 127.) But, IA has otherwise stated that the project was “led by Internet Archive’s legal counsel, Lila Bailey.” (*See* Exhibit 46 at INTARC00151102.) When the release of the White Paper was delayed, Bailey wrote to the group to say that “[t]he Internet Archive has more than an academic interest in this moving forward” and that “it will be deemed a complete failure and waste of our time if it’s not released by [an upcoming IA event]. I’m just being real about the impact of delay on my client (and potentially on my job if this isn’t done... seriously guys).” (A true and correct copy of the relevant email from Lila Bailey is annexed hereto as Exhibit 128.)

84. Other scholars were consulted during the project for their thoughts on the draft Statement, including Peter Jaszi – a copyright expert and advisor to the Internet Archive who attended the Workshop. (*See* Exhibit 46 at INTARC00151184.) (A true and correct copy of an email copying Jaszi on an email from Internet Archive to “Open Libraries Advisors & Supporters” is attached hereto as Exhibit 129.) An IA blog post refers to Jaszi as one of Bailey’s “heroes.” (A true and correct copy of the relevant blog post is annexed hereto as Exhibit 130.) Jaszi wrote to at least one librarian to state that the draft Statement was a “one-sided puff piece that seriously misrepresents both the state of the law and the risks to institutions of pursuing the strategy” and that

controlled digital lending “serve[s] the institutional interests of the IA.” (A true and correct copy of the relevant email from Peter Jaszi is annexed hereto as Exhibit 131.)

85. The library copyright expert Kenneth Crews also identified a number of problems with the draft Statement. (A true and correct copy of the relevant email from Crews is annexed hereto as Exhibit 132.) For instance, he wrote that “CDL use is NOT identical to current physical lending.” And he noted that if physical copies of books “are still available for non-circulating use, one could argue that you have doubled the number of readable and useable copies.” (*Id.*)

86. The Statement and White Paper on controlled digital lending were published simultaneously in September 2018. The Statement lists as co-authors Bailey, Courtney, Hansen, Minow, Shultz and Wu. (A true and correct copy of the Statement is annexed hereto as Exhibit 133.) The Statement does not disclose that Michelle Wu had been engaged as an attorney for Internet Archive from 2017 until at least March of 2018. (Bailey Tr. 82:9-17; 90:14-24.) The Statement asserts that “[p]roperly implemented, CDL enables a library to circulate a digitized title in place of a physical one in a controlled manner. Under this approach, a library may only loan simultaneously the number of copies that it has legitimately acquired, usually through purchase or donation. For example, if a library owns three copies of a title and digitizes one copy, it may use CDL to circulate one digital copy and two print, or three digital copies, or two digital copies and one print; in all cases, it could only circulate the same number of copies that it owned before digitization. *Essentially, CDL must maintain an ‘owned to loaned’ ratio.* Circulation in any format is controlled so that only one user can use any given copy at a time, for a limited time. Further, CDL systems generally employ appropriate technical measures to prevent users from retaining a permanent copy or distributing additional copies.” (Statement at INTARC00458737 (emphasis added).)

87. The Statement further claims that “there are two main areas of copyright law that support CDL: the principle of exhaustion and the fair use doctrine.” (*Id.*) The primary thesis is that the “first sale” doctrine – which allows libraries to distribute lawfully acquired physical books – allows creating and lending digital files of the physical books because “any time there is a lawful transfer of a copy of a copyrighted work, the rights holder’s power to control the use and distribution of that copy is terminated or ‘exhausted.’” (*Id.*) The Statement further asserts, in the absence of a “directly analogous case on point,” that “fair use and copyright exhaustion can work together to effectuate CDL practices.” (*Id.*) In essence, the Statement argues that controlled digital lending is not unlawful because the “purposes of library lending ... all focus on socially beneficial outcomes that favor fair use” and because there is no market harm “because properly implemented CDL programs maintain an ‘owned to loaned ratio’ that is comparable to physical lending.” (*Id.* at INTARC00458737-38.)

88. The White Paper provides additional legal analysis and detailed advice on how controlled digital lending should be implemented. The White Paper notes weaknesses to the controlled digital lending theory. For instance, it acknowledges that there is a “considerable point of concern that CDL is not clearly transformative.” (Exhibit 123 at INTARC00358262.) The White Paper also acknowledges that in “mass digitization cases involving books – *Google Books* ..., for example – courts have largely focused on how those projects enabled transformative access to information by enabling text search, as well as research uses, such as text and data mining.” (*Id.*)

89. In terms of advice for practitioners of controlled digital lending, the White Paper stresses, with italics, that “libraries must truly exercise *control* in the process.” (White Paper at INTARC00358246.) Thus, steps must be taken to ensure that “[w]hen the digital copy is being

read by a patron, ... the corresponding physical copy is restricted and unavailable for consultation” – for instance, by “rapidly removing [physical] books from open circulation” when the ebook copy is lent. (*Id.* at INTARC00358245, INTARC00358280.) Libraries must also “ensure that original works are acquired lawfully,” “lend each digital version only to a single user at a time just as a physical copy would be loaned,” “limit the time period for each lend” and “use digital rights management to prevent wholesale copying and redistribution.” (*Id.* at INTARC00358246.)

90. The White Paper also sets forth a number of “risk mitigation” measures that controlled digital lenders can implement to further reduce legal risk, including:

- a. adding artificial “friction” by extending the time between digital lends, more closely mirroring how physical books are lent and returned;
- b. “introduce characteristics that mimic physical degradation” (*e.g.*, each ebook can only be circulated a limited number of times;”
- c. limit the audience to “particular communities of users” (*e.g.*, students of an academic institution or “local residents”) “to limit the overall reach of the copy and therefore the potential market effect;”
- d. only distribute older works published pre-1989 and public domain books;
- e. only distribute out-of-print or orphan works;
- f. only distribute highly factual books.

(Exhibit 123 at INTARC00358278-85.) Internet Archive’s Policy Counsel testified that Internet Archive does not practice any of these risk mitigation suggestions. (*See* Bailey Tr. 219:7-229:1.)

91. The Statement and the White Paper received significant criticism, including from the Association of American Publishers (“AAP”) and the Authors Guild. The AAP released a “Statement on Flawed Theory of ‘Controlled Digital Lending’” on February 4, 2019, which stated

that “AAP strongly disagrees with the analysis of the White Paper and its call to libraries to copy and transmit entire books to the public in disregard of the law. CDL not only rationalizes what would amount to systematic infringement, it denigrates the incentives that copyright law provides to authors and publishers to document, write, invest in, and disseminate literary works for the benefit of the public ecosystem.” (A true and correct copy of the AAP’s “Statement on Flawed Theory of ‘Controlled Digital Lending’” is annexed hereto as Exhibit 134.)

92. The Authors Guild released a statement on January 8, 2019 entitled “Controlled Digital Lending is Neither Controlled nor Legal.” (A true and correct copy of the Author’s Guild’s “Controlled Digital Lending is Neither Controlled not Legal” statement is annexed hereto as Exhibit 135.) In their statement, opposing CDL, the Authors Guild noted that Internet Archive had “started rejecting notices sent by Guild members asking for unauthorized digital copies of their books to be taken down, citing that it ‘operates consistently with the “Controlled Digital Lending [sic].”’ (*Id.*)

93. The Authors Guild statement explained that Internet Archive’s “threat to author income and the ebook market comes from two directions: 1) unauthorized scanning and e-lending of books that were previously published only in physical formats would usurp the market for creating new ebook versions; and 2) instead of purchasing library ebook licenses (which are more expensive than consumer editions for good reason), libraries would simply digitize the print book from their collection, depriving authors and publishers of important licensing income.” (*Id.*) “Needless to say,” the statement continued, “if Internet Archive’s plans to expand Open Library to all libraries are realized, it would eventually decimate the market for library ebooks, put a massive dent in the ebook market in general, and usurp authors’ rights to bring their older works back to the market.” (*Id.*)

94. IA’s Director of Open Libraries, Chris Freeland, mocked similar complaints from an English authors’ union. After Tom Blake, a librarian at the Boston Public Library, forwarded him emails from the union’s chief executive complaining that CDL “prevent[s] [authors from] making a living by licensing the content of [their] work,” Freeland wrote, “Nice that they included their real interest: ‘negotiate and purchase lending licenses from the copyright owners.’” (A true and correct copy of emails from Freeland to Boston Public Library Librarian Tom Blake are annexed hereto as Exhibit 136.)

95. In December 2018, the Second Circuit issued its decision in *Capitol Records, LLC v. ReDigi Inc.*, 910 F.3d 649, 659 (2d Cir. 2018), in which the Court, among other things, rejected the first sale doctrine as a basis for reproducing used music files, finding that “[u]nauthorized reproduction is not protected by” the first sale doctrine and that “the making of such reproductions is not a fair use.”

96. Jonathan Band, who is counsel for the American Library Association (“ALA”), submitted an amicus brief to the Second Circuit in *ReDigi* on behalf of Internet Archive, the ALA and others. (A true and correct copy of the Internet Archive’s amicus brief is annexed hereto as Exhibit 137.) The brief stated that, “[l]ike *ReDigi*, Open Library seeks to replicate in the digital realm the sorts of uses that are permissible with a physical copy of a work. . . .” (*Id.* at p. 20.) Band thus urged the court to hold that *ReDigi*’s practices were covered by the first sale doctrine or fair use “to spur investment and innovation in digital services by libraries” – including Internet Archive’s “platform that makes millions of books available in various formats, and with varying degrees of technical restrictions on copying and further distribution.” (*Id.* at p. 19-21.) The brief explained that “A fair use finding in this case would provide libraries with additional legal certainty to roll out innovative services such as the Internet Archive’s Open Library.” (*Id.* at p. 5.) After

the *ReDigi* decision came down, Band published an article stating that it “raises questions concerning the viability of Controlled Digital Lending ... by libraries,” which “must be carefully reevaluated in light of this decision.” (A true and correct copy of Band’s article entitled “The Implications of the ReDigi Decision for Libraries” is annexed hereto as Exhibit 138.) Band warned that, “the decision calls into question the theoretical underpinnings of CDL. Specifically, CDL relies on the fair use right to replicate the first sale right in the digital environment. Judge Leval’s decision, however, could be read to suggest that the objectives of the first sale right cannot guide the fair use analysis.” (*Id.*)

97. Despite these concerns expressed by its own counsel in the *ReDigi* matter, Internet Archive has not made significant changes to its practices in response to *ReDigi*.

98. While some public libraries and individuals became “Signatories to the Position Statement on Controlled Digital Lending by Libraries,” academic libraries were the majority of signatories. (A true and correct copy of the list of “Signatories to the Position Statement on Controlled Digital Lending by Libraries is annexed hereto as Exhibit 140.) Of the 9,000 public libraries in the United States, which tend to have widely circulating collections of general interest books and circulate high volumes of authorized ebooks, only a handful endorsed the Statement. (*See id.*) Many copyright experts, including Band, Crews and Jaszi, did not sign on to the Statement. (*Id.*) Nor did the American Library Association. (*Id.*) And the MIT Press, which allowed IA to digitize some of its backlist books, also declined to become a signatory because “being able to monetize digital sales of trade books and text books is currently a significant part of what allows us to survive as an [open access] friendly mission driven publisher.” (A true and correct copy of relevant emails from Amy Brand, Director of MIT Press, is annexed hereto as Exhibit 141.) (*See also* email from Hansen to Bailey, Wu, Minow, Courtney and Schultz

indicating that Brandon Butler, an early signatory of the statement, wished to remove his name, a true and correct copy of which is attached hereto as Exhibit 142.)

INTERNET ARCHIVE IMPLEMENTS AND MARKETS THE OPEN LIBRARY PROJECT

99. Starting in about 2019, Internet Archive marketed its Website and CDL to libraries in connection with their outreach concerning the Open Libraries project (*see, supra* at ¶¶67-76) by relying on the Statement and White Paper to provide assurances on the legality of the project. For instance:

- a. On November 28, 2018, Chris Freeland sent an email to Janet Snowhill, the System Librarian for Chemeketa Cooperative Regional Library Service, following up on a solicitation to participate in Open Libraries to inform her of the publication Statement and White Paper. (A true and correct copy of this email is attached hereto as Exhibit 143.)
- b. On March 3, 2019, Chris Freeland sent an email in response to a question from Kristy Draper, the Digital Project Manager at Houghton Mifflin Harcourt about the security of IA's ebook storage by informing her that "The team of copyright experts that authored the white paper on Controlled Digital Lending have just published an FAQ that addresses file security.") (A true and correct copy of these emails is attached hereto as Exhibit 144.)
- c. Internet Archive published a blog post on July 29, 2020 stating that "CDL is a respectful and secure way to bring the breadth of our library collections to digital learners... Publishers are seeking to shut this library down, claiming copyright law does not allow it. Our response is simple: Copyright law does

not stand in the way of libraries' rights to own books, to digitize their books, and to lend those books to patrons in a controlled way." (Exhibit 9.)

- d. In February 2019, Internet Archive convened a panel at the Boston Public Library "to seek to educate the public on the work of the Internet Archive and Open Libraries" and the event description stated that "[t]hrough controlled digital lending, libraries can make twentieth century scholarship available that is largely absent from their digital holdings in a way that respects the rights of authors and publishers." (A true and correct copy of the event description for the panel is annexed hereto as Exhibit 145.)
- e. A July 1, 2019 blogpost by Kahle stated that "[w]e believe that every library can transform itself into a digital library. If you own the physical book, you can choose to circulate a digital version instead." (Exhibit 64.)
- f. Another Internet Archive presentation entitled "How Controlled Digital Lending Works for Libraries" contained a slide stating that "Controlled digital lending is a longstanding widespread library practice that you can use to help your patrons access digitized books. The Internet Archive has more than 1.8M digitized books you can claim & offer your patrons today. Your library can participate by joining Open Library." (Exhibit 72.)
- g. A February 11, 2021 "mythbusting" panel that "dispelled myths about CDL" by emphasizing "the limited and controlled aspect of the practice" and describing fair use as a "superpower" that "allows libraries to responsibly lend materials, and experts say logically includes both print and digital works [sic]." (A true and correct copy of the February 11, 2021 blogpost entitled

“Mythbusting Controlled Digital Lending: Community Rallies to Fight Misinformation About the Library Practice” is annexed hereto as Exhibit 146.)

100. Internet Archive does not indemnify Partner Libraries for liability arising from its infringement. As Freeland wrote to a librarian at Duke University, Kahle “generally gets allergic to indemnification clauses...” (A true and correct copy of the relevant email from Freeland is annexed hereto as Exhibit 147.) At least four Partner Libraries have withdrawn from the Open Libraries project since this action was filed. (*See* Exhibit 114.)

101. Internet Archive confirms that the “most critical” principle of CDL is the owned-to-loaned principle. (Freeland Tr. 224:24-225:4.) To comply with the “owned-to-loan” principle, a “library may only loan simultaneously the number of copies that it has legitimately acquired.” (Statement at INTARC00458736.) But the overlap analysis conducted as part of the Open Libraries project enables libraries to “pool[] their physical collections in order to make more lendable copies of digital books available...” (Exhibit 100.) Or, as Freeland put it in an email to a librarian, “libraries put one copy in for our matched records (even if you have multiple copies), and those counts are added to a pool for a given book. When a patron checks out a book, they are checking out a copy from the pool, not an individual library; for a variety of reasons, including reader privacy, we don’t connect the circulation back to an individual library.” (A true and correct copy of emails between Freeland and Kevin French is annexed hereto as Exhibit 148.)

102. Internet Archive admits that a match under the overlap analysis will “increase the number of concurrent borrowers by one, independent of the number that [the Partner Library] ha[s] in the library.” (Kahle Tr. 204:21-22.) It further acknowledges that this means that if three Partner Libraries contributed a book to the concurrent lending count under an overlap analysis and one of those libraries, for instance the MIT Library, only had one physical copy, three MIT patrons would

be able to read the ebook on the Website at once. (*Id.* at 205:23-207:16.) As Freeland testified, the overlap analysis dictates that, “under controlled digital lending, if a hundred Partner Libraries possessed a copy of the same book, the Internet Archive would be able to lend a hundred copies of that book simultaneously.” (Freeland Tr. 65:2-15.) And “if there were a thousand libraries that had the same book and put them into controlled digital lending,” Internet Archive could lend one thousand copies regardless of how many physical copies it owned. (*Id.*) Freeland also conceded that the overlap analysis allows Partner Libraries to loan more copies than they individually own (Freeland Tr. 92:24-94:13) – which creates inevitable “more loaned than owned” scenarios.

103. With its Open Libraries project, Internet Archive has turned itself into a centralized hub for ebooks by dropping geographic limits, “pooling” books to increase concurrent lending caps and encouraging libraries to incorporate links to ebooks on the IA Website regardless of what is in that particular libraries actual holdings. Under the current system, “[a]ll users have equitable access to the materials [Internet Archive] lend[s]. When a library adds a copy into [the] lending counts, that book can be checked out by any user.” (A true and correct copy of emails between Chris Freeland and Ellen Finnie, Hear of Scholarly Communications & Collections Strategy for MIT Libraries, is attached hereto as Exhibit 149.) Even non-partner libraries can integrate links to Internet Archive ebooks into their websites. (A true and correct copy of a July 30, 2020 blog post on archive.org is attached hereto as Exhibit 150.) (*See id.*, “To be clear, you don’t have to join the program to access our books. Anyone can link to our books right now.”)

104. Librarians have identified concerns with these practices. A librarian at Fordham University, for instance, contacted Internet Archive in July 2020 about the possibility of “digitiz[ing] approximately 450 volumes ... to support CDL for distance learning.” (A true and correct copy of relevant emails between Freeland and Michael Weiss, Assistant Director of

Technical Services at Fordham University Libraries is annexed hereto as Exhibit 151.) The Fordham librarian specifically flagged that “[t]hese are largely or entirely in-copyright books, so the scans could not be made public, and will only be released to our users in proportion to the number of physical copies which we have embargoed.” (*Id.* at INTARC00445587.) In response, Freeland replied “[t]hat’s not how our implementation of CDL *currently* works – we lend to anyone with an Internet Archive library card, and there is no way to restrict usage of books put into CDL by particular users, so there’s no way for us to limit use to only your patrons.” (*Id.* at INTARC00445585.)

105. Similarly, Freeland received an inquiry from Wendy Knapp, the Deputy Director of Indiana State Library, asking whether its copies of a book, in this case *Slaughterhouse Five*, “would be added to the total number of copies in OpenLibrary.org” or whether those books would start “a new waitlist just for Indiana patrons...” (A true and correct copy of the email exchange between Freeland and Knapp is annexed hereto as Exhibit 152.) Freeland responded that the Indiana State Library’s copies “would decrease the waitlist and be loaned to all users” of IA’s website. (*Id.*) Indiana’s Deputy Director responded that her Library “still have hesitation around the idea that a format shift has not been indisputably set as a fair use” and declined to join as an Open Libraries Partner Library. (*Id.*)

106. Internet Archive also admits that it exercises no actual controls over Partner Libraries that join the Open Libraries project, even though CDL instructs that “libraries must truly exercise *control* in the process.” (See White Paper at INTARC00358246.) Specifically, IA admits that it does nothing to require that the Partner Libraries who contribute books to the concurrent lending count ensure that the physical copies remain locked down while the ebooks circulate. The Open Libraries Form, for instance, contains no provisions on how libraries implement CDL

because Internet Archive considers this to be a “local decision.” (A true and correct copy of the Open Libraries Form is attached hereto as Exhibit 153.) (*See id.* at INTARC00142704.) In response to an email from a librarian expressing “concern[] about being in compliance,” Freeland wrote that “how libraries limit circulation for books in CDL is a local decision made by the library... [Some] libraries are taking the approach that they don’t need to suppress circulation because the likelihood is slim that all our digital copies and all the physical copies across our network of partners are all checked out at the same time.” (*See* Exhibit 148.) Thus, as Freeland testified, Internet Archive “was aware ... that some partner libraries did not suppress circulation when they agreed to bec[o]me a partner library and put their books into controlled digital lending.” (*See id.*)

107. Internet Archive also acknowledges that it has “no way of knowing whether a book ... was being read in a particular library at any given time” or “whether the physical and digital copies of the book were in circulation simultaneously.” (Freeland Tr. 174:3-175.) Nor does it know how Partner Libraries store the physical books counted in the overlap analysis. (*Id.* at 114:15-25.) Freeland is also aware that even if a library puts a physical book into a non-circulating reference collection, it could be read in the library while the ebook equivalent is checked out. (*Id.* at 171:6-172:18.) There is also no technological system by which Internet Archive informs Partner Libraries when an ebook in its collection is checked out, or to tell Internet Archive when the physical book is circulating. (*Id.* at 166:23-167:2.) And Internet Archive has never “taken any action against a library that did not suppress circulation properly.” (*Id.* at 174:24-175:3). IA’s founder has testified that “monitoring” libraries would amount to “hav[ing] people going and snooping around,” which it will not do. (Kahle Tr. 173:4-11.)

108. According to the Open Libraries Form, Internet Archive requires Partner Libraries to submit their catalogues on a “quarterly” basis so that Internet Archive can re-run the overlap analysis to periodically add new books that the library acquired and subtract books that dropped out of the collection as a result of weeding. (*See* Exhibit 153.) In fact, IA admits that in the first years of the Open Library project it did not “require updates at a specific time frame” and “sometimes Internet Archive did annual updates” depending on “the library’s preferences.” (Freeland Tr. 106:5-19.) Internet Archive also does nothing to ensure the accuracy of the catalogues it receives, but rather “depends on our partners for verifying the data.” (Freeland Tr. 62:11-17.) While IA began to run the overlap analysis on a monthly basis after this action was filed (Freeland Tr. at 105:22-24; 107:24-108:4), it still does not require Partner Libraries to inform it when they weed out books and no longer physically own the titles. (*See* Freeland Tr. at 108:15-109:3.) This means that concurrent lending counts for an ebook could routinely exceed the number of Partner Libraries with that physical book in their collection. (*Id.*) Internet Archive also does not “audit library collections to ensure that the library still owns the physical copy it’s lending against” because, as Kahle put it, the “Internet Archive doesn’t snoop around dumpsters or things.” (Kahle Tr. 176:7-11.)

109. There are also flaws in the overlap analysis that lead to inflated concurrent loan caps due to certain inaccurate or not properly organized data and metadata (*See* Foster Decl. ¶ 86.) As a blog post inviting users to “Volunteer @ Open Library” put it, “Open Library’s book catalog has millions of books and thousands of data errors. Sometimes author names are misspelled, book covers are missing, or works and authors are duplicated or conflated.” (A true and accurate copy of the blog post is annexed hereto as Exhibit 154.)

**INTERNET ARCHIVE’S VARIOUS AND SHIFTING POLICY RATIONALES FOR ITS
ACTIVITIES ARE NOT SUPPORTED BY THE EVIDENCE**

110. When seeking to garner support, participation from libraries, or funding, Internet Archive has claimed that CDL advances numerous important public interests – such as digitizing older books that do not exist as authorized ebooks, catering to the visually impaired, serving rural communities and enabling datamining – but these interests are not reflected on the homepage of openlibrary.org., which reads that the Website is to enable anybody in the world to “Read Free Library Ebooks Online,” including the “[m]illions of books available through Controlled Digital Lending.” (See Exhibit 12.)

111. Internet Archive asserted in the first paragraph of the Executive Summary to its application for the \$100 million McArthur Foundation grant that “there’s almost a century of knowledge still living only on the printed page, missing from our digital shelves.” (Exhibit 46 at INTARC00151088.) In 2018, Kahle wrote to the Librarian of Congress, Carla Hayden, stating that controlled digital lending was “a way to get 20th century books onto the net respectfully.” (A true and correct copy of a relevant email dated September 27, 2018 from Kahle to Hayden is annexed hereto as Exhibit 155.) In that email, Kahle also forwarded a blog post by Courtney and Hansen to the Librarian of Congress, Carla Hayden, stating that the goal of controlled digital lending is “particularly to help address access to the large number of books published in the ‘20th Century black hole’ that have little hope of otherwise being made available to readers” (*Id.*) More specifically, the “black hole” refers to the gap between the date when books enter the public domain (currently 1926) and “the 1990s.” (See Exhibit 64 at p. 2.)

112. The evidence shows, however, that Internet Archive actively seeks to post popular and widely read books. Kahle stated in “Transforming Our Libraries into Digital Libraries paper, that “[t]he Internet Archive, working with library partners, proposes bringing millions of books online, through purchase or digitization, *starting with the books most widely held and used in*

libraries and classrooms.” (See Exhibit 101 (emphasis added).) He continues, “[f]or the books we can not [sic] buy in electronic form, I am proposing a collaborative effort to select and digitize the most useful books of the 20th *and 21st centuries.*” (*Id.* (emphasis added).)

113. Internet Archive’s focus on popular books and genres is reflected in the homepage for its Website, which lists books in categories like “Trending Books,” “Romance,” “Thrillers,” “Kids,” “Classic Books” and “Books We Love.” (See Exhibit 12) There is no category on the homepage for “Lost 20th Century Books” or anything similar. (*Id.*) The “program lead for OpenLibrary.org,” Michael “Mek” Karpeles, confirmed that “there are tens of thousands of readers on Open Library who are looking for romance novels, self-help books, kids books, and modern thrillers.” (A true and correct copy of the relevant email from Karpeles to Cheng is annexed hereto as Exhibit 156) (A true and correct copy of relevant portions of the transcript of the October 27, 2021 deposition of Michael “Mek” Karpeles (“Karpeles Tr.”) is annexed hereto as Exhibit 157.) (See *id.* at 242:6-242:18.)

114. Nor does the data support IA’s suggestion that its Website primarily stocks mid-20th Century titles. The breakdown of ebooks by publication date on the Website shows that 78.9% of the in-copyright books were published after 1980 and only 6.2% of those books were published before 1963. (A true and correct copy of a spreadsheet of publication year data from archive.org/details/inlibrary is annexed hereto as Exhibit 158.)

115. Internet Archive has represented that it “limits its digital lending to books published in the past five or more years.” (A true and correct copy of IA’s pre-motion letter for its summary judgment motion in this case is annexed hereto as Exhibit 159.) But, two of the Works in Suit – *All the Presidents’ Women* and *The Man Who Solved the Market* – were published in 2019 and republished on the IA’s Website that same year. (See Foster Decl. ¶ 67.) Further, the remaining

Works in Suit were all published more than five years before the action was filed and continue to be significant sources of revenue for their authors and the Publishers. And OverDrive’s evidence documents that the sales of a particular work can spike years after the book is first published. As but one example, the book *Crazy Rich Asians* saw an enormous spike in popularity – including increased demand for library ebooks – when a major motion picture version was released five years after the book was first published. (A true and correct copy of Exhibit 13 to the Deposition of Steve Potash, reflecting OverDrive checkout data, is attached hereto as Exhibit 160.)

116. Another example is the book “The 5 Simple Fixes That Will Make You Healthy, Fit, and Eternally Awesome,” by Darin Olien, which was published by an imprint of HarperCollins in February 2015. For the first five years after it was published, the title had been checked out by library patrons across the United States in the low double-digits (and sometimes in the single digits) each month. (*See id.*) Then, in July 2020, checkouts spiked – eventually resulting in nearly 700 checkouts in September 2020 alone. The increased interest in the book – over five years after its publication – coincided with the July 2020 release of the Netflix series “Down to Earth with Zac Efron,” in which the famous actor Zac Efron would promote the “5 Simply Fixes” book at the beginning of each episode. (A true and correct copy of a July 10, 2020 *Men’s Health* article titled “Darin Olin is More than Zac Efron’s Travel Partner in Netflix’s *Down to Earth*” is annexed hereto as Exhibit 161.)

117. The Internet Archive also has represented to the Librarian of Congress that controlled digital lending is “not meant to be a competitor to *Overdrive*, nor a replacement for licensing e-books of best-sellers or other currently licensable e-book content.” (A true and correct copy of a September 27, 2018 email from Kahle to Carla Hayden of the Library of Congress is attached hereto as Exhibit 162.) This representation is not consistent with the representations that

IA made to libraries that participation in the Open Libraries project means “free ebooks for your patrons” with “no cost involved.” (*See* Exhibit 104; *see also* ¶71 *supra*.) Freeland has also testified that Internet Archive “add[s] ebooks” to the Website that are “available to purchase commercially or to license to libraries.” (Freeland Tr. 85:17-25.) Internet Archive does not do anything “to make sure that the books they add to the Open Libraries projects are not yet available in digital form” and does not “instruct libraries to segregate out books that are otherwise available in digital form.” (*Id.* at 85:3-9.) The 127 Works in Suit – which are all available as authorized library ebooks – were republished as unauthorized ebooks on the Website, together with 33,000 other titles available from the Publishers in digital formats. (*See* Foster Decl. ¶¶ 110-115.)

118. Any claim that Internet Archive is supposed to be different from aggregators like OverDrive is undermined further by the striking similarity between the two platforms. The home page of openlibrary.org strongly resembles the interface that public library patrons use to read authorized ebooks via OverDrive. (A true and correct copy of the OverDrive homepage for the Detroit and Austin Public Libraries is annexed hereto as Exhibit 163.) (*Compare* Exhibit 12, Open Library homepage). Both websites display rows of thumbnail book covers available to be “borrowed,” organized into categories of general interest such as “Trending” books. (*Id.*)

119. Internet Archive has acknowledged that its Website competes with distributors of the Publisher’s authorized ebooks. Karpeles wrote an email stating – and confirmed at his deposition – that “reality has it that we are competing for eyeballs with Amazon, Goodreads (45M monthly users), Overdrive, OCLC Worldcat, and countless other websites...” (*See* Karpeles Tr. 19:16-21; 240:24-242:5.) Karpeles also created a document listing the library ebook aggregators OverDrive and Hoopla, as well as the ebook retail platforms Amazon and Google Books – all of which make available the Publishers’ full list of works – as “similar services” to Internet Archive.

(A true and correct copy of the “Open Library Design Ecosystem” document is annexed hereto as Exhibit 164.) (*See id.* at pp. 2-3.)

120. Internet Archive’s witnesses also admit that the ebooks on the Website can be used as substitutes for authorized ebooks. For instance, Internet Archive’s library expert, Susan Hildreth testified that it is her “expert opinion that libraries will spend less money on licensing the e-book editions of the particular titles that were provided through CDL.” (Hildreth Tr. 226:24-4.) Hildreth further testified that “I think it could be likely that with CDL materials meeting some patron requests, the result that [sic] the library would not necessarily have to license the specific title to make patron demand, they would be able and it is likely that they would purchase additional e-materials ... to meet other patron demand.” (*Id.* at 43:8-45:5.) Hildreth acknowledged “that a specific author would not receive a royalty for a specific title” if revenue was shifted in this way. (*Id.*) Finally, Hildreth admitted that “certainly some [ebooks]” on the Website “could be read in an hour.” (*Id.* at 254:17-255:1)

121. Internet Archive users view the Website as a substitute for authorized library ebooks. A user of Internet Archive testified that he read an ebook on IA’s Website that he could have read in an authorized ebook format through his university library because “it didn’t matter to me ... whether I was using Internet Archive or [the authorized platform]... [I]t didn’t ... matter to me. Or it didn’t factor into my thinking because I was more concerned with getting these materials I needed...” (A true and correct copy of relevant portions of the transcript for the April 6, 2022 deposition of Daniel Smith (“Smith Tr.”) is annexed hereto as Exhibit 165; *see* Smith Tr. 73:4-77:16.) Another Internet Archive user who writes blogs distributing links to ebooks on Internet Archive testified IA’s Website is one of “lots of ways to acquire books” but she would not incorporate links to authorized library ebooks to her bibliographies because their distribution is

limited to members of particular public libraries, whereas Internet Archive is accessible to anybody in the world. (Gibbs Tr. 112:25-116:12; 122:11-123:22.)

122. Internet Archive also emphasizes that the ebooks on its Website serve the needs of print disabled users and stated in the Executive Summary of the MacArthur Foundation application that “[w]orking with US libraries and Benetech, operator of the world’s largest digital library for people with disabilities that impact reading, the Internet Archive (IA) will bring millions of free digital ebooks to billions of people. For the blind, ebooks are a lifeline...” (See Exhibit 46 at INTARC00151088.) But the vast majority of Internet Archive users are not print disabled. Only 2 print disabled users were added in 2017 and, by 2021, there were only 10,891 such users – which is less than 0.2% of the accounts currently registered with the Website. (See Exhibit 114, Response to Interrogatory 16.) Moreover, the needs of print disabled readers are already served by HathiTrust, which “provides print-disabled patrons with versions of all of the [20 million +] works contained in its digital archive in formats accessible to them” – and, unlike Internet Archive, HathiTrust truly focuses on the print disabled by prohibiting the general population from accessing full ebooks. *Authors Guild, Inc., v. HathiTrust*, 755 F.3d 87, 101 (2d Cir. 2014).

123. Internet Archive also stated in the Executive Summary of their application for the MacArthur Foundation grant that by “digitizing millions of books, we unlock them for communities with limited or no access” – particularly “people in rural areas.” (See Exhibit 46 at INTARC00151088.) But Internet Archive does not target particular geographic locations because “someone could sign up from anywhere in the world.” (Freeland Tr. 177:7-12.) Internet Archive also has acknowledged that it would not know whether a user checking out an ebook from the Website “was coming from New York City or coming from some small town in Upstate New

York” because Internet Archive only maintains “state- or country-level” geographic data about users. (*Id.* at 117:5-25.)

124. When commenting on the Statement, the library copyright expert Kenneth Crews noted that the argument that CDL is particularly necessary to serve individuals in rural areas is “not very convincing. Many rural communities are well served. Sometime[s] the biggest need is in the cities, where parking is miserable, and realistically students and profs are doing their research at home.” (*See* Exhibit 132.) The founder of OverDrive testified that “many of the public libraries that are [his] customers ... serve rural communities,” “serve disabled communities,” “serve poor communities or impoverished communities” – including by OverDrive making donations when appropriate. (*See* Potash Tr. 172:23-173:12.) National data indicates that rural communities have seen more growth and value from authorized library ebooks than libraries serving more densely populated districts. (A true and correct copy of the Institute of Museum and Library Services report entitled “The Use and Cost of Public Library Materials” is annexed hereto as Exhibit 166.)

125. Further, in many rural states, the state-wide or regional library systems pool their resources to collectively license ebooks that are made “available to all residents of that state, whether urban or rural” and “creates greater availability to all types of potential readers, whether they be rural or disabled or poor or whatever.” (Potash Tr. 191:6-9, 20-23.) Many rural libraries make use of this opportunity to lend on a state-wide basis or as part of regional consortia. (A true and correct copy of the webpages reflecting the North Dakota Digital Consortium, Mississippi eBook Library Partnership, Oklahoma Virtual Library, and the Idaho Digital Consortium) is annexed hereto as Exhibit 167.)

126. Internet Archive also has indicated in its MacArthur Foundation application that the ebooks on its Website can be used for data or text mining because “[a] digital repository of more than four million books and related metadata creates a dataset of great interest to researchers and data scientists.” (Exhibit 46 at INTARC00151100; Answer at p. 2). But Freeland testified in his capacity as Internet Archive’s 30(b)(6) witness on this topic that he does not know “what percentage of users engage” in those practices and knows of only three projects that could be described as datamining. (Freeland Tr. 204:10-208:25.)

THE NATIONAL EMERGENCY LIBRARY

127. On March 24, 2020, Internet Archive launched a “National Emergency Library” in the wake of COVID-19 pandemic shutdowns of libraries. (A true and correct copy of the March 24, 2020 blogpost announcing the National Emergency Library is annexed hereto as Exhibit 168.) The National Emergency Library drew from the same pool of ebooks that were previously available on the Website, but it “suppressed the wait-list function” – *i.e.*, “[d]id away with the one-to-one ratio.” (Kahle Tr. 182:11-22.)

128. In practical terms, IA “did away with waitlists” and “did away with the one-to-one ratio” by raising the cap on concurrent loans to 10,000. (Freeland Tr. 223:5-20.) Demand for ebooks on the Internet Archive spiked after lending limits were relaxed and, even after those limits were reimposed, key metrics like monthly circulation and number of total members have remained higher than their pre-pandemic levels. For example, as of September 2019, IA reported over 2.7 million total “Members” of the Website; by April 2022, that figure had nearly doubled to over 5.5 million. (A true and correct copy of captures from the Open Library Stats page is annexed hereto as Exhibit 169.) Likewise, as of March 2020, IA reported that the number of “ebooks borrowed” “over the last 28 days” on the Website was 41,296; by July 2022, the number of ebooks borrowed over the last 28 days is listed as over 349,984 – a more than eightfold increase. (A true and correct

copy of captures from the Open Library home page reflecting this data is annexed hereto as Exhibit 170.)

129. Since Internet Archive did not require users “to certify that they were directly impacted by COVID-19” to access IA’s ebooks on demand and without waitlists, anybody in the world with an internet connection could access the National Emergency Library. (Freeland Tr. 242:11-18.) Bailey wrote that this was because “we don’t have the time or staffing to allow us to limit NEL access only to people directly impacted by COVID 19.” (A true and correct copy of Bailey’s email containing this statement is annexed hereto as Exhibit 171.) And while Internet Archive offered rightsholders an option to “opt-out” by sending takedown requests, books that “were taken out of the National Emergency Library, ... were not taken out of controlled digital lending.” (Freeland Tr. 250:12-20.)

130. Internet Archive was aware that there were “copyright concerns” about the National Emergency Library. (Freeland Tr. 221:15-223:4.) Senator Tom Udall of New Mexico wrote to Maria Strong, the U.S. Register of Copyrights, to urge her “to examine the National Emergency Library that has been organized by the Internet Archive which is operating without typical library license and is causing authors in New Mexico concern about the integrity of their copyrights.” (A true and correct copy of Sen. Udall’s letter and Register Strong’s response is annexed hereto as Exhibit 172.) Senator Udall noted that “[s]ince the emergence of e-books, libraries have provided e-books to readers through legally well-established means of paying for licensing fees for e-books that they lend, of which a portion of the licensing fees extend to authors as royalties.” (*Id.*) But given the National Emergency Libraries distribution of ebooks without payment, he had “heard from authors who are concerned that such action is not legal and presents additional challenges to them at an economically difficult time.” (*Id.*)

131. In response, Register Strong explained that despite the Internet Archive’s claim “that the books in the National Emergency Library ‘focus on materials published during the 20th century, the vast majority of which do not have a commercially available ebook,’ and which therefore would not be publicly available when schools and libraries are closed[,]” “the Internet Archive does not appear to have verified if any of the works in its collection were available to the public in digital formats prior to including those books in its collection or removing its waiting lists.” (*Id.*)

132. In her letter, Register Strong also explained that “[t]he types of materials included in the National Emergency Library, which include Stephen King thrillers and joke books, also suggest that at least some of the materials are likely to be accessed for entertainment rather than educational purposes” and that, as a general matter, “[t]he Copyright Office has also consistently expressed doubt that providing digital access to complete works can be considered a fair use.” (*Id.*)

133. Two days after the launch of the National Emergency Library, Alan Harvey, the Director of Stanford University Press made “a formal request to exclude all [its] content from the National Emergency library.” (A true and correct copy of Harvey’s email is annexed hereto as Exhibit 173.) He explained that “[w]e do not grant those rights to our content, and do not agree that the current coronavirus emergency constitutes an argument for copyright violation. To support the community during this crisis, we are in the process of making all our content more liberally available through the existing library aggregators, and responding to individual requests for content access.” (*Id.*) In a subsequent email to Freeland, the Stanford University Press Director stated Internet Archive’s practices were “entirely unacceptable” given that “IA simply grabbed everything without any discussion and the bad will is now oozing through us all.” (*Id.*)

134. On March 31, 2020, the Director the University of Minnesota Press wrote an email to Freeland stating that “the National Emergency Library goes further than we can legally or ethically allow.” (A true and correct copy of this email is annexed hereto as Exhibit 174.) After noting that takedown requests had been sent, the Director stated that he would “be happy to discuss making limited Minnesota content available as part of the National Emergency Library, understanding that we **always** consult authors or rights holders before making such decisions – indeed, we’ve just done a round of that for course books we’ve made openly available for emergency course use on our Manifold OA platform. We’d also need some kind of binding legal agreement.” (*Id.*)

135. On April 2, 2020, the Executive Director of the Authors Guild, Mary E. Rasenberger, wrote to Kahle to “set up a time to talk” about the National Emergency. (A true and correct copy of Rasenberger’s email is annexed hereto as Exhibit 175.) She wanted to “dispel any rumors you have heard that the Authors Guild is planning to bring suit... We are instead addressing this bold infringement by letting authors know how they can request IA to take their books down, and we hope to appeal to your decency in asking you to shut this ill-planned (even if well-intentioned) initiative down. There are plenty of other sources for books for students and teachers right now. We can help you point people to them.” (*Id.*)

136. Kahle’s response: “We talked before, and I felt deposed by a lawyer... I talked with multiple of your board members with no apparent effect. We are actively working with others and are making great progress. Wish us all luck, we are all trying to get a digital world that works for everyone.” (*Id.*) Internet Archive also ignored public complaints by the AAP. (*See* Exhibit 134.)

137. On April 8, 2020, PRH approved a request by a school for special COVID-19 class set pricing for *Interview with the Vampire* by Anne Rice and *Lolita* by Vladimir Nabokov. (A true and correct copy of the emails indicating PRH’s approval of the request is annexed hereto as Exhibit 176.) But PRH was informed by OverDrive that “the school will not be moving forward with a purchase. They cited that they needed the materials more urgently and the professor was able to find both titles on the Internet Archive / DPLA site with simultaneous checkout so they used those versions.” (*Id.*)

138. On April 14, 2020, Kahle wrote an email to the Executive Director of the Authors Alliance and Pamela Samuelson (who sits on its Board of Directors) to notify them that “the Authors Guild is going to be meeting with some staffers [in Congress] about copyright matters, and possibly about NEL as well. I thought I would alert you to this, as the Authors Alliance, I think, was formed to be at the table when these discussions were happening.” (A true and correct copy of Kahle’s email is annexed hereto as Exhibit 177.) Kahle himself is on the Advisory Board to the Authors Alliance and Dave Hansen – one of the White Paper’s co-authors – is on the Staff. (A true and correct copy of the webpage for the Author’s Alliance Advisory Board is annexed hereto as Exhibit 178.) Internet Archive asked the Authors Alliance to endorse the National Emergency Library but it ultimately did not “endorse the National Emergency Library.” (A true and correct copy of an email from Brianna Schofield to Bailey, Samuelson and Freeland is annexed hereto as Exhibit 179.) (*See also* Freeland Tr. 255:5-24.)

139. Kahle also posted a blog post on April 14, 2020 entitled “The National Emergency Library – Who Needs It? Who Reads It? Lessons from the First Two Weeks.” (A true and correct copy of the blog post is annexed hereto as Exhibit 180.) He wrote that “[W]e moved in ‘Internet Time’ and the speed and swiftness of our solution surprised some and caught others off guard. In

our rush to help we didn't engage with the creator community and the ecosystem in which their works are made and published. We hear your concerns and we've taken action: the Internet Archive has added staff to our Patron Services team and we are responding quickly to the incoming requests to take books out of the National Emergency Library. While we can't go back in time, we can move forward with more information and insight based on data the National Emergency Library has generated thus far." (*Id.*)

140. Internet Archive ended the National Emergency Library on June 10, 2020, ten days after the Publishers filed this action and two weeks before the earliest possible date IA had previously announced for its closure. (A true and correct copy of the blog post announcing the end of the National Emergency Library is annexed hereto as Exhibit 181.) Kahle announced that Internet Archive would return to "traditional controlled digital lending" pursuant to overlap analysis, which it continues to practice to this day. (*Id.*)

141. On September 4, 2020, the Executive Director of HathiTrust, Mike Furlough, declined an invitation from Kahle to speak on a panel called "Digital Lending in a Pandemic." (A true and correct copy of Furlough's email is annexed hereto as Exhibit 182.) During COVID-19, HathiTrust expanded access to its ebooks in the wake of school library closures but imposed "many limitations" that IA did not, including restrictions on access to "students, faculty and staff as designated by the affected institution" and "access ratios in proportion to the number of physical copies already owned by a given library that are temporarily inaccessible." (A true and correct copy of Skip Dye's email summarizing HathiTrust's and IA's approaches is annexed hereto as Exhibit 183.) (*See also* the terms for HathiTrust's Emergency Temporary Access Service a true and correct copy of which is annexed hereto as Exhibit 184.)

142. Furlough declined the invitation to speak because “[i]f I am asked to define the difference between what we have done and what you have done and are now doing, I will end up pointing out that we have put a good many more *controls* on the service that you did for NEL or for ongoing lending ... I may have to end up directly contrasting some things, which would implicitly suggest a criticism of your approach. And we made some very different decisions – when we did our legal analysis, it pointed us in directions that are quite different from those you have taken. And we did not feel that the methods you were employing were ones *we* could defend. We didn’t think we could rely on the same legal theories that you were relying on. I’d rather keep all that to ourselves and not be quoted or paraphrased as suggesting that your actions are NOT defensible, especially at your own party.” (See Exhibit 182.)

I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct and that this declaration was executed on July 7, 2022.

/s/ Elizabeth A. McNamara
ELIZABETH A. MCNAMARA